



MATHEMATICAL ECOLOGY

OF POPULATIONS AND ECOSYSTEMS

JOHN PASTOR

VISIT...

LANZAROTE
Caliente.COM

VISIT...

LANZAROTE
Caliente.COM

Mathematical Ecology of Populations and Ecosystems

To
my grandparents, István and Erzsébet Szajkó,
my parents, Joseph and Mary Pastor,
my wife, Mary Dragich,
and my son, Andrew

Mathematical Ecology of Populations and Ecosystems

John Pastor

*Professor, Department of Biology
University of Minnesota Duluth
Duluth, Minnesota
USA*

 **WILEY-BLACKWELL**

A John Wiley & Sons, Ltd., Publication

This edition first published 2008, © 2008 by John Pastor

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical and Medical business to form Wiley-Blackwell.

Registered office: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloguing-in-Publication Data

Pastor, John.

Mathematical ecology of populations and ecosystems / John Pastor.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-4051-7795-5 (pbk. : alk. paper) – ISBN 978-1-4051-8811-1 (hardcover : alk. paper)

1. Ecology–Mathematical models. 2. Ecology–Mathematics. 3. Population

biology–Mathematical models. I. Title.

ISBN: 978-1-4051-8811-1 (plpc) and 978-1-4051-7795-5 (pb)

A catalogue record for this book is available from the British Library.

Set in 9.5/12pt Berkeley by Graphicraft Limited, Hong Kong

Printed in Singapore by COS Printers Pte Ltd

Contents

Prologue	vii
Preface	ix
Acknowledgments	xiii
Part 1: Preliminaries	1
1 What is mathematical ecology and why should we do it?	3
2 Mathematical toolbox	11
Part 2: Populations	51
3 Homogeneous populations: exponential and geometric growth and decay	53
4 Age- and stage-structured linear models: relaxing the assumption of population homogeneity	65
5 Nonlinear models of single populations: the continuous time logistic model	78
6 Discrete logistic growth, oscillations, and chaos	92
7 Harvesting and the logistic model	110
8 Predators and their prey	129
9 Competition between two species, mutualism, and species invasions	159
10 Multispecies community and food web models	176
Part 3: Ecosystems	187
11 Inorganic resources, mass balance, resource uptake, and resource use efficiency	189
12 Litter return, nutrient cycling, and ecosystem stability	218
13 Consumer regulation of nutrient cycling	238
14 Stoichiometry and linked element cycles	252
Part 4: Populations and ecosystems in space and time	271
15 Transitions between populations and states in landscapes	273

16	Diffusion, advection, the spread of populations and resources, and the emergence of spatial patterns	284
	Appendix: MatLab commands for equilibrium and stability analysis of multi-compartment models by solving the Jacobian and its eigenvalues	298
	References	305
	Index	319

Prologue



Here is a photograph of a forest in northern Sweden. It is a pine forest near my office in the Department of Animal Ecology at the Swedish University of Agricultural Sciences in Umeå, Sweden, where I wrote much of this book while on sabbatical during 2005–2006. I would often take a break from writing about mathematical ecology and walk through this forest to refresh my contact with the natural world, as a stimulating contrast with the abstract world of this book.

The forest is used to teach forestry students about the measurement and management of such lands. Occasionally, parts are thinned, and the evidence of repeated thinnings and other management activities are readily apparent. Clearly, the large older trees were spared thinning to provide a seed source for the next cohort of pines.

None of the trees are “original old growth” but, according to my colleagues, there has always been a forest here, perhaps since the Vikings. There is a Viking burial ground on one of the knolls, looking out over what must have once been a bay of the Baltic Sea when the land was lower – it has risen by almost three meters since Linneaus’s time because of rebound from the glaciers. Obviously, the forest had seen what mathematical ecologists dryly refer to as “perturbations” but equally obviously it had recovered and persisted in some recognizable form.

Questions about the persistence of the forest and recovery from repeated perturbations over the centuries suggested themselves. How is the current number of live trees of different ages related to the numbers of seedlings established decades and centuries ago? How do the decay of the annual cohorts of dead needles and other debris of the forest floor replenish and recycle nutrients taken up by the plants from the soil? What about the animals which wander through here, such as the moose who left pellets (more debris!) behind and browsed the pines? How do they affect the dynamics of the plant populations and the cycling of nutrients?

For that matter, what do we mean by all these abstract terms: *perturbation*, *persistence*, *dynamics*, *decay*, *growth*, *populations*, *cycles*? I could not point to any of these concepts like I could point to a pine tree or a moose or the soil, but I could not think about the forest without using them. How can we think clearly about how these abstract terms relate to the plants, animals, and soil?

That is what this book is all about.

Preface

This is an introductory textbook on mathematical ecology bridging the subdisciplines of population ecology and ecosystem ecology. The expected reader is you: a beginning graduate student, advanced undergraduate student, or someone who thinks of themselves as a student all their lives, with a working knowledge of basic calculus and basic ecology. While this is intended as a stand-alone text, the level is such that once you have read through it, you will be able to read more advanced texts and monographs such as Ågren and Bosatta (1998) and Kot (2001) with greater depth. While there are other very good introductory texts in mathematical ecology (e.g., Edelstein-Keshet 1988 [reissued but not revised 2005], Yodzis 1989 [now out of print], Gotelli 1995, Roughgarden 1997, Case 2000, and Kot 2001 are among the most widely used), none bridge the gap between population ecology and ecosystem ecology.

Ecological problems are complicated in ways that our language has not evolved to handle. Mathematics provides a more precise way than the spoken word of thinking and talking about the rates of change, nonlinearities, and feedbacks characteristic of populations and ecosystems. Even Edward Abbey, one of the sharpest thinkers who never solved an equation, once said: “Language makes a mighty loose net with which to go fishing for simple facts, when facts are infinite.”* It is my hope that mathematics can help us tighten that net a bit, allowing us to catch a few facts that may have otherwise slipped through. At the very least, mathematics has precise tools for handling the infinite.

For the past 12 years, I have given courses in mathematical ecology and ecosystems ecology during alternate fall semesters. Often, the major topics of our discussions in each of these courses concern the relationships between population dynamics, species, and ecosystem processes such as productivity, nutrient cycling rates, and input–output budgets. These are leading research questions in ecology and have been major interests of mine for the past 25 years.

These are intellectually challenging questions. It is often easy to make a “plausible” argument that some hypothesized relationship between populations, species, and ecosystems must be true, only to find on more rigorous examination that it is not necessarily true, true only under certain restrictions, or simply not true at all. Framing the plausibility argument in mathematical terms and using the rules of mathematics to examine its logical structure is often the best way to uncover the sense in which it might be true. In fact, the mathematical examination of these arguments often uncovers hidden assumptions; these in turn suggests new experiments to determine

**Desert Solitaire*, Author’s Introduction.

whether they hold in the “real” world or new theoretical investigations to determine what happens when such hidden assumptions are relaxed in different ways.

Questions about the relations between populations and ecosystems are also challenging because the cultures of population ecologists and ecosystem ecologists differ so much. In part, these differences between population and ecosystem ecologists arise in their graduate training. Population ecologists are often trained to analyze the dynamics of populations’ and species’ interactions using analytical mathematical methods which allow them to calculate algebraic expressions for equilibria and their stability.

Because population models focus on the dynamics of collections of live individuals, death is often treated as an export from the system. By contrast, the ecosystem ecologist considers dead material to still be in the system, simply be detached from the live populations and subject to different rules. Eventually, through microbial decay, the dead material is transferred to the resource pool which is then taken up by plants.

Because of the large number of compartments they generally consider and measure, ecosystem ecologists have not usually used analytical mathematical methods. Instead, large and complicated computer simulation models have traditionally been their method of choice in analyzing and synthesizing ecosystem data. Although these simulation models may make quite accurate predictions for specific situations, they are often almost as complex as the system being studied. Therefore, it is sometimes difficult to understand why their predictions are as they are, leading to an interesting paradox in which accurate prediction may not be the same as general understanding. Furthermore, through their investigations into the origin of chaos in single-species models, population ecologists have taught us that we can have understanding without predictability, a conclusion accepted with reluctance by some (but not all) ecosystem ecologists.

Therefore, ecosystem ecologists and population ecologists have been trained to speak different languages. Population ecologists have traditionally ignored nutrient feedbacks to populations through litter and its decay, whereas ecosystem ecologists traditionally dismiss analytical approaches in favor of simulation models. This lack of a common language or approach amongst population and ecosystem ecologists may impede our ability to address important practical problems. It is no wonder that many ecologists find the relationship between populations, species, and ecosystem properties extremely difficult to understand: each group has part of the answer but they find it difficult to speak to each other and frame questions in a common language.

In spite of the traditional dichotomy of using either simple analytical models of a few species or complicated simulation models of whole ecosystems, it is possible to couple interactions between species and the flux of an inorganic resource by simplifying the ecosystem to only a few compartments so that we can use analytical mathematical techniques to gain understanding about system behavior, especially how ecosystem properties emerge from an interaction between populations and the flux of inorganic resources. In this book, the same mathematical techniques will be used as a common thread to help unify population and ecosystem ecology.

These mathematical techniques are also essential for exploring how changes in controlling factors across thresholds often cause rapid changes between different states of an ecological system, which are sometimes called “regime shifts” (Scheffer et al. 2001). These rapid changes between different states are often accompanied by the

appearance of new behaviors, such as limit cycles, extinction of species, or changes in top-down and bottom-up controls. Some examples of current interest are the possibility of rapid change in communities and ecosystems with slowly rising temperatures once some critical value of temperature is exceeded, the rapid changes in communities once critical thresholds of nutrient inputs are exceeded, and the rapid changes and extinctions in populations once critical values of harvesting rates are exceeded.

Different states of populations, communities, or ecosystems often correspond to different equilibrium solutions of a model. In turn, these solutions are often separated by a critical value of a parameter or a function of several parameters. Rapid changes in the nature and stability of solutions of equations as critical parameter values are crossed are known mathematically as bifurcations. Bifurcations between different equilibrial solutions appear suspiciously like the rapid changes in nature as controlling factors cross critical thresholds. Examples of bifurcations which we will meet in this book include: (i) saddle-node bifurcations, separating persistence from extinction of a species once a critical harvesting rate is exceeded; (ii) transcritical bifurcations, leading to shifts between two different communities once critical inputs of a limiting nutrient are exceeded; (iii) Hopf bifurcations, leading to stable limit cycles once critical values of carrying capacity are exceeded, otherwise known as the “paradox of enrichment” (Rosenzweig 1971); and (iv) Turing bifurcations, leading to the appearance of spatial patterns once critical values of diffusion rates of populations are exceeded. We shall explore examples of these and other bifurcations and their ecological implications throughout this book.

Bifurcation theory is therefore a powerful mathematical technique to help us understand sudden and interesting changes in the behaviors of ecological systems as some parameter or combination of parameters pass some critical value. Bifurcation theory draws heavily on the theory of eigenvalues and Jacobians and, insofar as bifurcation theory seems a promising mathematical approach to understand rapid changes in nature, one must have some grounding in eigenvalue analyses – indeed one must be able to frame questions and construct systems of equations with the use of these techniques in mind.

The purpose of this textbook is therefore to help you develop your thinking to bridge population and ecosystem problems using the mathematical tools of eigenvalue analysis and bifurcation theory as common threads. To successfully do this, you need a working understanding of calculus, especially the concept of limits; linear algebra, especially matrix operations required to analyze populations with age or stage structure or multiple species models; and differential and difference equations, especially the analysis of model stability by means of eigenvalues and eigenvectors. While all ecology graduate students have had training in calculus, it may have been a while since they used it; a few have had experience of linear and matrix algebra; very few have been exposed to eigenvalues and eigenvectors. Accordingly, Chapter 2 is a “mathematical toolbox” laying out the tools to be used in this book and providing some exercises for you to practice using these tools without much reference to any biology at first. This lays the foundations for a mathematical vocabulary for the book. Many of these exercises will appear later in more ecological form.

I try whenever possible to derive the standard equations of mathematical ecology from some more fundamental “first principles” of birth and death, probability of two individuals meeting, and conservation of matter. Typically, these derivations are

motivated by uncovering or relaxing some “hidden assumption” to address an unrealistic behavior in some prior, simpler model. In addition, intermediate steps in these derivations often shed some light on what the final equation means: a lack of understanding of where the final equation came from can lead to misleading analogies and conclusions. In addition, many derivations and proofs often depend on some trick or turn of an argument in an intermediate step, and learning these tricks or turns of an argument both enriches the ecological and mathematical underpinnings of a model and often proves useful in derivations of other models.

Every chapter begins with an introduction to a new problem, usually motivated by some problems unearthed in the previous chapter or chapters. These problems are usually an unrealistic biological behavior of the previous, simpler models. We then try to uncover the assumptions that may be responsible for the problem behaviors. The chapters usually proceed by mathematically relaxing these assumptions in different ways and analyzing how this improves the model’s behavior (or not).

Every chapter ends with two sections, the first entitled: “Summary: what have we learned?” which, besides the obvious summarizing of the main points, also brings the discussion back to a wider plane. The final concluding section of each chapter (except the Introduction and Mathematical Toolbox) is a section called “Open questions and loose ends.” Here, I point you in some directions and towards some papers or texts about problems that lack of space does not allow me to go into. I also suggest some open questions for you to consider. Some of these are small questions for you to explore, perhaps as additional homework problems, but they may lead to larger questions. Some of these are large open questions (such as control of chaos in population models) which are at the current edge of research. I hope that these may help you choose a thesis problem (if you are a graduate student) or research problem (if you are already establishing your own program). I would welcome learning from you any findings along these lines or about any papers that have addressed them that I may not know about (and for which I apologize to the authors).

By introducing you to the mathematical tools required to analyze models of populations, communities, and ecosystems, I hope to help you develop more rigorous ways of thinking about the interaction of population and ecosystem dynamics. It is my further hope that these ways of thinking will spawn more creative approaches to these problems.

I have learned much by writing this book: oftentimes, connections have emerged that neither I nor (I believe) anyone else has seen before. If you are already a professional mathematical ecologist or mathematician, I hope that these connections will surprise you as much as they did me. If you are a student, I hope you will learn as much or more than I did and, in turn, teach me through the papers you will write.

Acknowledgments

This book grows out of a graduate course in Mathematical Ecology which I have taught for the past twelve years, in both the Biology Department at the University of Minnesota Duluth and in the Department of Animal Ecology at the Swedish University of Agricultural Sciences in Umeå, Sweden. Teaching this course is always one of the high points of my year. I must therefore first thank all the students who have taken this course over these years. They have helped me clarify and simplify various explanations of mathematical ecology in my lectures and I hope some of their help has worked its way into this book. During 2006, students in both Sweden and Minnesota read through drafts of these chapters during class and I thank them in particular for catching typographical and other errors, for pointing out where more explanation is required, and for suggesting simplifications of some explanations and derivations.

Most of these chapters were written during 2005–2006, while I was on sabbatical leave in Umeå. Financial support for this leave came from the College of Science and Engineering at the University of Minnesota Duluth, the Department of Animal Ecology at the Swedish University of Agricultural Sciences, and the Kempe Foundation, and I thank them all for their generosity.

I am especially grateful to my colleague Kjell Danell of the Department of Animal Ecology at the Swedish University of Agricultural Sciences for helping to arrange my sabbatical visit and the grant from the Kempe Foundation. Through Kjell's help, I was provided with a quiet office with a view of a forest where I could write and think about mathematics and ecology, and my wife Mary and I were provided with an excellent apartment from which we could ski off into the forest right from our door. Kjell, his family Kerstin Huss-Danell and Markus Danell, and my colleagues at the Department of Animal Ecology provided superb hospitality in the best Swedish tradition, and to all of them I say: *Tack så mycket!*

Special mention must be made of my colleagues Bruce Peckham and Harlan Stech of the Department of Mathematics and Statistics at the University of Minnesota Duluth and Yossi Cohen of the Department of Fisheries and Wildlife at the University of Minnesota St. Paul. I have collaborated with them over the years on topics both mathematical and ecological. I have learned much from each of them, and I hope the things I have learned from them show in this book. Bruce Peckham helped clarify my thinking on several of the topics and Harlan Stech read through the entire book and made many helpful comments and suggestions and corrected some errors. Tom Andersen of the University of Oslo also read many of these chapters and offered helpful comments and encouraging words. I thank Harlan, Bruce, and Tom for their help. Any remaining errors remain my own and I ask that if you spot one, please notify me of it.

Parts of several of these chapters were presented at the weekly seminar of the Department of Mathematics and Statistics at the University of Minnesota Duluth. I thank the faculty and students at these seminars for their insights and helpful comments.

Alan Crowden guided the proposal for this book through the review process and presented it to Blackwell Publishing. Without his encouragement to begin writing and his help and assistance with the publishing world, this book may not have been begun at all. Ward Cooper of Blackwell also provided publishing assistance, and I appreciate his efforts and those of his staff, especially Rosie Hayden, Pat Croucher, and Delia Sandford, as well.

Rachel MaKarrall scanned and lettered the figures; their clarity owes much to her artistic eye.

Two anonymous reviewers took the time and care to read through the manuscript and made many helpful comments and suggestions. I thank you both and hope you find the revised manuscript improved as a result of your efforts.

But above all, I must thank Mary Dragich, my wife, who has always given me support and encouragement, especially during the writing of this book, and who has listened patiently and helpfully to my long explanations of ecological and mathematical problems over the dinner table.

John Pastor

Part 1

Preliminaries

1 What is mathematical ecology and why should we do it?

Let's begin by looking again at the photograph in the Prologue and imagine yourself walking through this forest. What do you see? Jot down a few things (this is your first exercise). They need not be profound – in fact, it is best not to try to make them profound. After all, Darwin constructed the most profound theory in biology by asking ordinary questions about barnacles, birds, and tortoises, amongst many other things.

Perhaps you see big trees and little trees and think that big trees are older than little trees. You also might notice that there are more little trees than big trees, and so not every little tree grows up to be a big tree – most die young. But the little trees must come from somewhere, namely seeds produced and shed by the bigger trees. These are the core ideas of population ecology.

Or perhaps you might notice that there are some dead needles and leaves on the ground and some standing dead trees which will eventually fall to the soil, the result of the deaths of those young trees and plant parts. You also note that the live trees have roots in the soil formed partly from those dead leaves and logs and surmise that the trees obtain some nutrients from them. These are the core ideas of ecosystem ecology.

These two views of the forest look very different, but they both contain biological objects that interact with each other through hypothesized processes. When we model a biological object such as a population, we begin by offering an analogy between it and a mathematical object. Mathematically we will term these analogs *state variables*. The processes usually represent a transfer of something (live individuals, seeds, nutrients) from one biological object to another. Processes will be modeled by mathematical *operations*, such as addition, multiplication, subtraction, or powers. One or more operations and the objects they operate on will be encapsulated into an *equation*, specifically an equation which relates how one state variable partly determines the state of itself and perhaps another at some point in the future. These equations will contain, besides mathematical operations and state variables, some *parameters*, whose values remain fixed while the state variables change. Each state variable will be described by one equation. The time-dependent behavior of the state variables and the magnitudes of the state variables at equilibrium are called the *time-varying* and *equilibrium solutions* of the model, respectively. We then use the rigor of mathematics to work through the logic of our thinking to gain some insight into the biological objects and processes.

Therefore, mathematical ecology does not deal directly with natural objects. Instead, it deals with the mathematical objects and operations we offer as analogs of nature and natural processes. These mathematical models do not contain all information about nature that we may know, but only what we think are the most pertinent

for the problem at hand. In mathematical modeling, we have abstracted nature into simpler form so that we have some chance of understanding it. Mathematical ecology helps us understand the logic of our thinking about nature to help us avoid making plausible arguments that may not be true or only true under certain restrictions. It helps us avoid wishful thinking about how we would like nature to be in favor of rigorous thinking about how nature might actually work.

What equations should we choose to use to model the dynamical relations amongst the state variables? Of course, there are an infinite number of equations we can choose, but we prefer equations that are simple to understand, are derived from simple “first principles,” have parameters and operations that correspond to some real biological process and are therefore potentially measurable, and produce surprising results that lead to new observations. These four properties of these equations are components of mathematical beauty. They are important criteria by which we judge the utility of an equation or model because they help clarify our thinking. They often force our thinking into new directions.

This is all well and good, but why should we play this game? Why not just state hypotheses as clearly as we can and do the experiments to test them? One reason is that we are often not sure of either the internal logic of our ideas and hypotheses or their consequences. For example, state variables often affect and are affected by another state variable. This mutual interaction between state variables is termed *feedback*. Feedbacks are common in ecological systems – in fact, they are characteristic of all interesting ecological systems. Systems with internal feedbacks are almost impossible to completely understand in an intuitive way. Without a clear understanding of how the feedback works, it is also very difficult to do an experiment which manipulates the feedback. It is easy to understand a chain of events where X influences Y and Y influences Z , but what if Z also affects X ? What then happens to Y ? By writing a system of equations, one for each of the state variables and using the rules of mathematics, we can examine the logical structure of feedbacks and their consequences.

Examining the properties of a system of equations allows us to pose further questions and determine how their answers might follow logically from their structure and properties. For example, the population ecologist might wonder how the proportion of individuals of a given age class changes over time, whether the proportional distribution over all age classes ever settles down to a stable distribution, and what that distribution is. The ecosystem ecologist might note that the world surrounding the forest contributes material to it (in rainwater, for example) and the forest contributes material back to the surrounding world (in the water leaching out of the soil). He or she might wonder what difference it makes how and where the material enters and leaves the ecosystem. Both ecologists might also wonder what happens if we harvest some of a population or ecosystem: does the population or ecosystem recover to its earlier state? How will it recover? Can we harvest so much that the population or ecosystem will never recover? And what exactly is meant by “recover”?

Examining these equations also allows us to uncover hidden assumptions about our ideas and ask what happens when we relax those assumptions. For example, we have assumed that each equation in our model applies equally well to every species that is reasonably similar to the one we are studying. Well, do they? What difference does it make if they aren’t similar to each other? How different do things have to be

to make a difference in the system's behavior? How do different species affect each other? How does including additional trophic levels or other components affect the behavior of the models?

Finally, mathematical modeling allows us to rigorously connect the two different views of population and ecosystem ecologists. For example, the ecosystem ecologist notices that the forest floor contains layers corresponding to different ages of leaf litter from many years in the past. One year's leaf litter is transferred into older decay classes with each passing year. If the leaves are decomposing, something is being lost from each age class of litter. The ecosystem ecologist pauses and notices that these ideas bear a great deal of resemblance to the age class model of the population ecologist. Can we take the equations for the dynamics of the live populations and extend them belowground into the leaf litter? This shows the real power of mathematical abstraction. Once you recognize a structural correspondence between two different systems, then the same equations and same mathematical techniques could apply to both. If it turns out that this is the case, then the ecologist has discovered some underlying principle of organization in nature, a principle which he or she did not expect when first observing a particular forest (or prairie or lake) and jotting down what first caught his or her eye.

And that is what mathematical ecology is about.

The nature of theoretical problems and their relation to experiment

In the process of abstracting nature into a mathematical model, we run into a number of theoretical problems. These are distinct from the sorts of problems experimenters have to deal with. Most ecologists are familiar with experimental questions such as measuring the response of an individual, population, or ecosystem to manipulations, or determining the proper number of samples required to detect a difference between mean values of measurements. In contrast to these experimental problems, mathematical models of ecological systems address a variety of theoretical questions regarding the logical consistency and consequences of ideas (Caswell 1988). While measuring devices are the tools of the experimental ecologist, equations are the tools of the mathematical ecologist. Equations are used to examine the following theoretical problems (Caswell 1988):

Exploring the possible ranges of behavior of a natural system. In order to understand why a particular natural system behaves as it does, it is useful to discover the range of behaviors that is possible for the system to exhibit. The behavior of a particular natural system is simply one realization of a family of possible behaviors. Models delimit the theoretical range of behaviors that follow from simple assumptions (mass balance, birth and death, etc). Experiments delimit the actual range of behaviors realized in nature, or the realized subset of the set of possible behaviors. Sometimes, by delimiting the full range of possible behaviors, models indicate new areas where experiments need to be performed that no one had previously realized, such as in extreme environments.

Exploring the logical consistency of ideas with a set of common axioms. Upon detailed examination, we often find that many plausible ideas are not consistent with some simple assumptions we must make about nature. Mathematical models allow one to logically connect an idea or a hypothesis with some axiom about nature. Reiniers (1986), for example, offers several axioms upon which ecosystem ecology might be based. Often, such theoretical exercises show that our hypotheses may be simply wishful thinking. It is often said that beautiful theories are killed by ugly facts, but it is equally

true that a beautiful hypothesis can be killed by being inconsistent with some more fundamental axiom of how nature works.

Exploring the connections between different ideas or experimental results by deriving them from a common set of assumptions. Oftentimes in ecology, different camps take up one side of an argument or another, resulting in “either–or” false dichotomies. Ecology is rife with these “either–or” arguments: either competition is important, or it is not; food webs are controlled either by top-down forces or by bottom-up forces; etc. The key word in all these arguments that creates the problem is the conjunction “or.” Usually, there is ample experimental evidence for both sides of the argument and so it is impossible for experimental ecology to clearly decide on one side or the other. At such times, it is useful to ask: when does one thing happen and when does the other thing happen? It may well be that there is some common underlying model that produces both sides of the argument at different time scales, for different parameter values, or for different initial conditions. Finding such a model and showing the conditions that lead to one system behavior or the other is a very important theoretical problem.

Evaluating the robustness of different approaches. An experimental result may be consistent with a particular way of simplifying nature, but how robust are our conclusions to uncertainties in the details of the structure of the natural system? Do we need to represent every age class in a population model, or can we aggregate age classes? Do we need to measure the population dynamics of every microbe to predict the fate of a nutrient during decomposition, or can we aggregate microbes into “microbial biomass”? How precisely do we need to measure minute-by-minute changes in photosynthesis to predict tree growth several years into the future? Can we even make predictions far into the future or is the natural system inherently sensitive to very small differences in initial conditions? How fast does the accuracy of our predictions decay with time?

Finding the simplest model capable of generating an observed pattern in nature. Such a model would suggest the simplest set of experimental protocols required to experimentally characterize a natural system. It could also pinpoint exactly which processes, interactions, or parameter values are responsible for observed behavior. Whether such a model is true to reality remains to be tested by experiment.

Predicting critical (falsifiable) consequences of verbal or conceptual theories. Prediction is considered to be a precise numerical value for something that can be measured, and so it often is. But prediction can also be qualitative, such as the shape of a response curve. The shape of a response can distinguish one mechanism from another. For example, different theories of nutrient uptake may yield response curves with different shapes, suggesting that experimenters test hypotheses about mechanisms of nutrient uptake by distinguishing between uptake curves of different shapes (O'Neill et al. 1989). Prediction can also be as simple (and as powerful) as postulating the existence of a particular behavior, such as the existence of limit cycles or other forms of complex population dynamics (Turchin 2003) or a decline in nutrient use efficiency at low levels of nutrient availability (Pastor and Bridgman 1999). At an early stage of experimental investigation, precise prediction of the magnitude of response may be unnecessary and being overly concerned with precise prediction or “validation” may even obscure broader issues of which mechanism is actually operating. We will have more to say about prediction and its role in model evaluation shortly.

Errors of perception of mathematical models

There are several common errors of perception of mathematical models (Caswell 1988):

The only thing to do with a theory is to test its predictions with experiments. This has to be done, but this ignores the role that rigorous mathematics can play in helping us work out the logic of our ideas before we even begin to design or execute an experiment. Much effort has been spent by myself and others on collecting data that in the end bears no relationship to the hypothesis being tested. Sometimes, this is fine because it helps us put the experiment in a larger context. It also allows us to serendipitously make connections between processes that might otherwise not have been made. But, since it also takes time and effort to collect data which may turn out to be unnecessary, we may also miss collecting data that is essential. Every modeler has had the experience of an experimentalist friend showing up with a boatload of hard-won data and asking for help to construct a model, only to have to say upon examining the data that much of it is not relevant to the experimenter's own statement of their hypothesis or that some key data required to construct a model of the hypothesis was not collected. In the latter case, the modeler then says that we will have to assume certain values. The conversation then usually deteriorates. The point here is that we should know which data are essential to the test of a hypothesis and which are ancillary, albeit desirable for other reasons. When we translate a hypothesis into a mathematical model, the attempt to precisely define each parameter and variable in terms of an analogous biological process or object helps clarify the essential data we need to collect.

Theories that are refuted by experiment should be abandoned. They can also be modified. Perhaps the experiment is in error or itself has ignored an important process. Data themselves may be in error, perhaps because of an unrecognized sampling bias. We should be as skeptical of data as we are of theories. As Sir Arthur Eddington once said, "Do not believe an experiment unless it is confirmed by theory."

Modelers make assumptions, which are evil, and the worst assumption is that the system is simple. Like many models, every experiment is based on a set of hidden assumptions. For example, the statistical analysis of experimental data makes the assumption that the natural system can be explained by linear models even though the system being manipulated is clearly non-linear. Models at least make assumptions explicit and also explicitly show the logical consequences of those assumptions, while the assumptions of an experiment often go unrecognized. In addition, uncovering and relaxing hidden assumptions of a model is a powerful theoretical tool to advance our understanding, one that we shall use throughout this book.

The simplicity of many models often brings out strong reactions from many experimenters, who are often upset when a process that they have spent their career studying and which is clearly operating in nature is not included in a model. The model is then often said to "oversimplify" nature and should therefore not be trusted. This is a healthy skepticism but it could also be directed against the experiments themselves. For example, most experiments (including my own) manipulate only two or three factors and measure the response of a single state variable, while many models consider two or more state variables and more than two or three parameters. Therefore, many experiments often simplify the natural system of interest even more than models.

What do we expect of mathematical models?

Some ecologists (e.g., Peters 1991) have argued that ecologists should concentrate solely on making quantitative predictions from models. Such a recommendation has much to recommend it, not least of which is that it will facilitate the interface between

models and experiments by demonstrating very specific and falsifiable consequences of hypotheses.

Prediction is a good thing when you can get it, but we cannot always get it. Prediction is often regarded as the highest test of a scientific theory – indeed, the ability to quantitatively predict something is often taken as the hallmark of a “hard” science. The epitomes of the hard, predictive sciences are, of course, physics and astronomy. However, predictive capability or the lack thereof may be “. . . an essential difference between the biological as against the physical sciences, rather than a *sine qua non* of scientific synthesis as such” (Holton 1978). There are several reasons why quantitative predictions are easier in physics than in ecology.

First, the entities that much of physics deals with, such as electrons and other particles, are identical in all pertinent respects, whereas the basic entities of ecology, namely individual organisms, vary quite a bit in their pertinent properties (and necessarily so, as Darwin taught us). Second, physical relationships are often linear. A linear model is one in which the size of something changes in proportion to itself or the size of something else (we will explore linearity in more rigor in Chapter 2). Linear models, as we shall see, exhibit simple behaviors. After all, much of the physics of the everyday world is derived from $F = ma$, which is a linear model. Ecological processes are inherently nonlinear (the size of something changes out of proportion to itself or the size of something else), and we shall see that nonlinear models exhibit very complex and surprising behaviors, stabilities, and instabilities (or bifurcations) with small changes in parameters near critical values. Predicting the behaviors of nonlinear systems using nonlinear models is a daunting task.

The statistical design and analysis of experiments is based on linear models of expected values of variables. We do not know how to design and analyze a nonlinear experiment. Instead, we experimentally break the system into linear chunks within which predictions are robust and easily falsifiable or verifiable. Models are useful in reassembling those chunks and synthesizing results of many experiments. Finally, it is well to note that in the branches of physics that deal with nonlinear processes, such as turbulence in fluid dynamics, prediction is every bit as difficult as in ecology.

Nonetheless, the physicist’s ability to simplify a problem to its essentials so that first and foremost the tools of mathematical rigor and logic can be brought to bear on the problem is a useful lesson for ecologists. When certain reasonable and almost axiomatic constraints – such as the conservation laws – are imposed on our equations, the number of solutions is minimized. The constraints on the model also tend to sharpen the differences between the solutions. This style of research produces insights of remarkable clarity. The solutions could represent different communities, for example, and discovering how the parameters and variables of a model lead to different solutions can give great insight into what controls the diversity of life without needing the model to make quantitative predictions. Predictive ability is sought only after a general mathematical analysis of the situation and the possible controlling factors. It is this style of thinking – the homing in on the essentials of a problem and its translation into mathematics – that I think ecologists can borrow from physicists, not the ability to make extremely precise predictions of properties, which may be something peculiar to much of physics.

Not being able to make a prediction should not prevent us from grappling with ideas. I don’t think Peters is saying that we should abandon an approach when it

cannot give a quantitative prediction. Rather, I believe he is saying that we should always have prediction in mind as a goal to work towards and in this I agree with him. But too much of an emphasis on quantitative prediction can blind us to the often more important and interesting qualitative behaviors of a model, such as when limit cycles or spatial patterns suddenly appear.

Another problem with concentrating solely or even mainly on prediction is that it can be quite easy to get very good predictions without gaining any understanding of nature. For example, one can obtain a very long time series of data (weather data or long-term population data, for example) and fit a model to it that is essentially a sum of sine functions of different amplitudes, frequencies, and phases. In principle, one can eventually get a model that goes through every point and will probably make accurate predictions into the future, for a while at least. But why do the sine waves have different frequencies, amplitudes, and phases, or for that matter why should sine waves describe the data at all? What are the sine waves trying to tell us about how nature works? Again, building models solely by fitting sine waves to data is a bit of a caricature and nobody is really suggesting that this and only this is what we should be doing, but it does serve to point out the problems of an overemphasis on predictability as the goal of modeling.

Finally, there is the class of models (used frequently in ecosystem ecology) known as simulation models, which are very complicated computer codes that try to depict processes explicitly and often give very good predictions. These are valuable tools and have certainly helped advance ecology and should not be abandoned. But, in my own experience with simulation models whose development I have been a part of, these models are often nearly as complicated as the system they are attempting to depict. Therefore, while it is wonderful to see trajectories of ecosystem development emerging on your computer screen when these models are run, why those trajectories are developing the way they are can be rather mysterious. There is a temptation (which I myself have felt) for simulation modelers, when asked: "How do you think such-and-such system works?," to hand the questioner a disk containing computer code and say "Just run this and you will see precisely what I think." Needless to say, the questioner does not always feel enlightened by this answer. Precise prediction of a wide variety of natural phenomena using a model that incorporates many conceivable ecological processes operating over a wide range of spatial and temporal response scales is impressive output, but we want something other than this from a model or theory.

What we require first and foremost of a model or theory is not prediction or reproduction of experimental results, but that it deepens and extends our understanding of nature, or at least our understanding of how we think about nature. By understanding I mean that the model transparently shows how various complicated phenomena, such as population cycles or sharp boundaries between ecosystems, emerge naturally from the basic ecological processes of birth, death, immigration and emigration, uptake of nutrients and water through roots or uptake of carbon dioxide and energy through leaves, and consumption of one species by another. How much of the complicated phenomena we see around us can be explained through these few processes? Transparency of assumptions and model structure, how the model relates to basic biological processes, and emergence of surprising results that bear some resemblance to complicated behaviors of natural systems are the main things we expect from models.

How mathematics helps us better understand nature is actually not well understood. Eugene Wigner called this “the unreasonable effectiveness of mathematics in the natural sciences” (Wigner 1960). Wigner asked why should some of the most abstract mathematical ideas of which we have no direct sensory experience have such an uncanny ability to describe the natural world and deepen our understanding of it? Neither Wigner nor anyone else has been able to answer that, but anyone who has experienced this consilience between mathematics and the natural world knows that it is a beautiful gift (Wilson 1998).

The mathematical models we will explore in this book have had a long history of deepening our understanding of the ecological world. Their simplicity makes some of the consequences of basic biological processes transparent but at the same time they exhibit behaviors that surprise us. As we think more deeply about why we are surprised by a model’s behavior and why it conforms to similar behaviors of real populations and ecosystems, we gain a deeper understanding of why certain things and not others might be happening in nature. Simplicity, transparency, emergence of surprising results, and understanding are what we seek. You must be the judge of whether you find them in any model.

2 Mathematical toolbox

In Chapter 1 we drew attention to the analogical, even metaphorical, relation between mathematical objects and operations and biological objects and processes. The use of metaphor and analogy in scientific reasoning is very deep and perhaps not as well appreciated as it should be (Holton 1986). Because analogies between two things are slippery, before we can properly use an analogy we must have a solid understanding of both sides of it. Otherwise, an analogy becomes “. . . a form of reasoning that is particularly liable to yield false conclusions from true premises” (Holton 1986). But if mathematical reasoning about the real world is an analogy, it has the peculiar strength of allowing us to determine exactly where our reasoning went false or whether we would arrive at the same conclusions from different but equally true premises. In order to do this, we must understand mathematical objects and operations *as mathematical objects and operations* first before we offer them as analogs of biological objects and processes.

In this chapter, we will develop and explore some basic mathematical objects and operations which will be used as analogs of biological objects and processes, respectively. We will use this material in the rest of the book to sharpen our thinking about ecological objects, the processes that they affect and which, in turn, act upon them. If you have a working familiarity with limits, matrices, and eigenvalues, you can skip this chapter. If not then it would be useful to spend time on this chapter so that later we can concentrate on the biological meaning of the derivations and analyses of these models without making too many side diversions into the mathematical techniques. I hope that this chapter will introduce you to some of the ways mathematicians think about things. I will occasionally “look ahead” and let you know the type of ecological problems for which this chapter’s material will be used. But for now, I mostly want you to just think about mathematics without worrying too much yet about how it will be applied to ecology. Let’s begin with numbers, which are perhaps the simplest of all mathematical objects, but as you shall see, they are deceptively simple.

Numbers, operations, and closure

The origin of mathematics may lie in the needs of primitive humans to characterize the magnitude of foods in relation to the size of a social unit: is there enough of something to feed all of us? Numbers might have been invented as convenient shorthand to keep track of essential resources and to determine if there is at least a 1:1 correspondence between the magnitude of resources and the size of the social unit among which the resources need to be shared (Barrow 1993). It is interesting to think that at least one plausible reason for the origin and early development of mathematics might be a problem that will concern us throughout this book, namely the relationship between a population and the resources that it must draw upon.

It is difficult to imagine how one could even begin to “think” about population density or the size of a pool of a particular nutrient without using numbers. Thinking about populations or nutrients seems to require using numbers at such a basic level that ecologists rarely give numbers and their mathematical properties a second thought. It may come as a surprise to many “pure” experimentalists that when they take data, they are implicitly doing mathematical modeling because they are associating a mathematical object (a number) with a property of a biological object. Indeed, we often proceed as if properties are defined in terms of numbers without realizing that numbers are mathematical objects used to model some perceived magnitude of a property. And so it may be good to begin our exploration of mathematical ecology by examining the properties of numbers and the basic operations which are performed upon them. Much of the discussion of numbers that follows is loosely based on chapter 22 from Feynman et al. (1963), which belongs in the repertoire of any literate scientist.

We need to develop a set of numbers upon which we can perform operations. We wish the set to be closed under the operation. By closure is meant that an operation on any element of a set of objects will produce another object that is also a member of that set. Closure avoids the problem of generating, by means of a mathematical operation analogous to a biological process, a mathematical object which does not belong to the set of mathematical objects that represents a set of biological objects. Without closure, the analogy between mathematical and biological objects and operations/processes would break down.

The natural numbers 1, 2, 3, . . . and 0 are the simplest sort of number. We will take these as given and not go into the set theoretic notions required to construct them from more basic concepts. What sort of operations are the natural numbers and zero closed under? Which biological processes are modeled by these operations and are these all the operations we need?

Beginning with 0, we generate the next natural number by adding 1:

$$a' = a + 1$$

Adding 1 to an object b times always generates a new natural number:

$$a' = a + b$$

And so natural numbers are closed under addition. Suppose we start with 0 and add a to it b times. We then get the operation of multiplication:

$$a' = 0 + a \times b$$

Since

$$a \times b = \underbrace{b + b + b + \cdots + b}_{a \text{ times}}$$

natural numbers are also closed under multiplication. If we start with 1 and multiply it by a b times in a row, we get the operation of powers, under which natural numbers are also closed:

$$a' = a^b$$

Besides the condition of equality, there is also the inequality conditions of $>$ (greater than) and $<$ (less than), which we will also find useful. Addition and multiplication of natural numbers are also closed under these inequality conditions, sometimes known as monotonic laws (Klein 1932):

$$\text{if } b > c, \text{ then } a + b > a + c$$

$$\text{if } b > c, \text{ then } a \times b > a \times c$$

Natural numbers are closed not only under addition, multiplication, taking powers, and the monotonic laws individually but also under various combinations and sequences of these operations. Combining these three operations give us a rich set of possibilities with which to manipulate natural numbers. Natural numbers are closed under the following sets of combinations of these operations:

$$a + b = b + a \qquad a + (b + c) = (a + b) + c$$

$$ab = ba \qquad a(b + c) = ab + ac$$

$$(ab)c = a(bc) \qquad (ab)^c = a^c b^c$$

$$a^b a^c = a^{(b+c)} \qquad (a^b)^c = a^{(bc)}$$

There are also identity operations which take a number and return it. The natural numbers are obviously closed under the identity operations. The identity operation for addition is adding zero, that for multiplication is multiplying by 1, and that for powers is raising any number to the power 1.

Functions are particular combinations of objects and operations. A function takes some mathematical object or objects, $x \in A$, performs operations on them, and returns exactly one other object $y \in B$. The objects, x , which functions take and change are called the arguments of the function, or the independent variables. Independent variables are usually acted upon by parameters, which are constants. The returned value, y , is the dependent variable. The set of values which the function takes, A , are called the *domain* and the set of acceptable values for the returned value, B , is called the *range* of the function. Although a function must return only one value y for each value of x , it can return the same value for several values of x . Functions are usually expressed as equations, but a table consisting of a column of values of an independent variable with a corresponding paired column of values of the dependent variable is also a function. A contour map is also a function that takes a latitude, longitude pair of numbers and returns a single elevation for that point on the surface of an object (such as the Earth). Whether they are tables, maps, equations or algorithms, functions associate each argument with a single returned value through their operations.

The standard symbol for the value of an argument is x (or, for many population models, N) and the returned value from the function is $f(x)$ (or $f(N)$). Strictly speaking, the function itself is not $f(x)$ – the function is f , which stands for the “rules” or sequence of operations. Sometimes we see the statement $y = f(x)$, but that means y is the returned value after the function operates on x ; y is not the function. For example, in the function $b = 5a$, a is the argument, 5 is a parameter, b is the returned value, and the function is the rule: “multiply a by 5 and return the value b .”

One of the keys to creating helpful biological models is to be able to associate each variable with a measurement of a biological object that can change in time or over space (e.g., mass, population density, nutrient content), each parameter with some constant property of the object (e.g., specific heat, input, per capita birth rate or death rate), and the operations with some biological process (e.g., birth, death, decay, harvesting, etc.). The relationship between a hypothesis about how a process operates on biological objects and the analogous mathematical function is then clarified.

For each operation, we also want to have an inverse operation that undoes the result of the first operation. The inverse of addition is subtraction. If $a + b = c$, then $a = c - b$. Now, the set of natural numbers is not closed under subtraction because subtraction of one natural number from another does not always yield another natural number. For example the operation $3 - 5$ makes no sense in the set of natural numbers. What should we do? Clearly, subtraction is sometimes very useful as the above example shows. Let's assume that subtraction is "true" in some sense and define new mathematical objects that are closed under subtraction so that we can use them to solve equations without generating something undefined. This is of course the set of integers, which is all the natural numbers plus negative whole numbers. This object allows us to solve the above equation, yielding $3 - 5 = -2$. Enlarging the set of objects to maintain closure under new operations is the strategy that mathematicians have used when faced with similar situations. This strategy has the advantage that the new set of objects remains closed under all the previous operations that the more restricted set is closed under and it is also closed under the new operation.

Using the set of integers, addition, and its inverse subtraction, we are now ready to make our first mathematical model of a biological process. This is a model for the population growth of a bee colony: There is one queen that produces b offspring over some as yet unspecified but still definite period of time and a workers who don't reproduce. The set of positive integers is the mathematical object of which one member is analogous to the size of the hive, b is another object from the set of positive integers analogous to the number of new offspring, and adding is the mathematical operation analogous to reproduction by one queen. N_0 is the size of the original colony of one queen and a workers before the queen reproduces and N_1 is the size of the colony after the queen produces b offspring. Thus, the new population is formed by the bees in a manner analogous to addition on natural numbers:

$$N_0 = a + 1$$

$$N_1 = N_0 + b = (a + 1) + b$$

If we subtract the first equation from the second we get:

$$N_1 - N_0 = (a + 1) + b - (a + 1) = b$$

and so b is the rate of change of the bee colony (number of new individuals) over some span of time.

Let's make a model of a different kind of population in which each of a organisms can produce b offspring apiece during some unspecified but still definite time period, such as a population of protozoan that can reproduce by mitosis or that of a plant species in which each individual can produce viable seeds. We now have a new

population which requires the operation of multiplication to stand for reproduction instead of addition:

$$N_0 = a$$

$$N_1 = N_0 + bN_0 = a + ba$$

Subtracting the first equation from the second gives the net change in the population during the time period, or

$$N_1 - N_0 = a + ba - a = ba$$

Thus, multiplication on the set of integers is a mathematical operation that yields a new model of population growth, one that is faster than addition because $a > 1$ and therefore $ba > b$. We will later show that this is the basic idea underlying exponential growth. Therefore, not only is the growth of a bee colony slower than that of populations of many other organisms, also it is fundamentally different because it is modeled by different mathematical operations.

These two examples are known as recursive models because they take the previous returned value of the dependent variable and use it in the next step as the argument to get or “map” the next value for the dependent variable. Recursive models operate in steps and the rate of change obtained by subtracting one step from the next is called a difference equation. There is an implicit delay in the process going from one step to the next. This delay in recursive models causes some peculiar behaviors, as we shall see in Chapter 6.

Integers thus allow us to count discrete objects such as individuals and manipulate them by means of multiplication, addition, and its inverse subtraction. But there are many quantities that interest us which can't be counted because they vary continuously, such as mass, population density per unit area, light levels, nutrient content, and nutrient concentration. These we must measure, not count. The measured values usually fall between the discrete values used for counting. Thus, to model measurements we must be able to express a fraction of an integer unit. This, of course, requires division. Division is also the inverse of multiplication but the integers are not closed under the operation of division: $3/5$ does not yield a number that can be expressed as an integer. Therefore, for the empirical reason of being able to express continuous measurements and also the mathematical reason of being able to use the operation of division, we have to proceed further by assuming division to be closed for all members of another class of numbers and define that class, which is the set of rational numbers (e.g., $3/5 = 0.600\ 000$). A feature of all rational numbers is that the sequence of numbers to the right of a decimal point eventually either falls into a repeatable pattern, or ends with $000 \dots$

It is easy to prove there are an infinite number of rational numbers. Take any rational number p/q in which p and q do not have any common divisors (otherwise we can simply factor them out). Add 1 to either p or q or both. The result is another rational number, p'/q' because p and q are both integers and integers are closed under the operation of adding 1. Generate yet another rational number by applying the same processes to p' and q' , *ad infinitum*.

But unlike the integers, rational numbers are “dense” in the sense that the gaps between successive integers are filled by the rational numbers such that between any

two rational numbers, no matter how close, we can always construct another. Here's how: take any two rational numbers p and q where $p > q$. Then there is some distance $p - q$ between p and q . Divide this distance by any number (2 will give the midpoint of the distance, so let's use 2 for convenience). Add this new number, $(p - q)/2$, to q . We now have a new number $q' = q + (p - q)/2$, which is also rational because it can be put into the form $(p + q)/2$. Therefore, the average of p and q is another rational number, q' , that lies between p and q (i.e., $q < q' < p$). Now construct another rational number between q and q' in the same way and iterate *ad infinitum*. We have just demonstrated how to construct a sequence of rational numbers that gets as close to q as we please and without ever reaching a stopping point by remaining in an interval from q to q' which is arbitrarily small. Note that we have not defined density as every rational number being "next to" or "immediately adjacent to" another rational number. Instead we have defined density as a sequence of operations which yields a rational number which can be made arbitrarily close to q by constructing intervals that are arbitrarily small but that still contain at least one rational number. This is a preview of problems that will be addressed by the concept of limits.

What happens when we take negative numbers as a power? In other words, what is the value of: $a^{(3-5)}$, where a is a natural number? Division provides the answer. We know from one of the above operations under which natural numbers are closed that

$$a^{(3-5)}a^5 = a^3$$

and by using the definition of division as the inverse of multiplication on a number,

$$a^{(3-5)} = a^3/a^5 = (a \times a \times a)/(a \times a \times a \times a \times a) = 1/(a \times a) = 1/a^2$$

but $a^{(3-5)}$ also equals a^{-2} , so $a^{-2} = 1/a^2$. Therefore, taking a negative power of a natural number is equal to the multiplicative inverse of a number raised to the positive magnitude of that power.

Division gives us the ability to more precisely and operationally define what we mean by "rate of change." In the above simple models of population growth, the rate of growth is somehow captured by b new individuals per reproductive individual over some unspecified but still definite period of time. How do we measure b ? Well, we go out and watch a single reproductive individual over some period of time and count the number of new individuals produced. But clearly the data we obtain will depend on how long we count – someone else could get a different estimate of b simply by counting for a different length of time. We don't want our measurements to depend on such trivialities as how long to count, so we need to express rate in terms of some basic unit of time. To do this, we must count not only the numbers of new individuals but the numbers of time units elapsed during the measurement. We then divide $N_1 - N_0$ by $t_1 - t_0$, where t_1 and t_0 are the end and beginning clock readings during the time of our counting:

$$\frac{N_1 - N_0}{t_1 - t_0} = \frac{(a + 1) + b - (a + 1)}{t_1 - t_0} = \frac{b}{t_1 - t_0} = b'$$

We now have a new estimate of the rate of change standardized to a time unit that everyone can agree upon. While this may seem to be a simple enough matter that any child in elementary school can be taught to perform, our ability to depend on this operation derives from the fact that we have defined a class of numbers – the rational numbers – that are closed under division (we glossed over the case that they are not strictly closed under division by all numbers because division by 0 is undefined. The way to deal with that problem is the development of the concept of limits and calculus, which we will come to in a bit.)

So taking inverses of operations seems to require us to expand our set of numbers and create a new set that is closed under all previous operations plus the new inverse operation. What about power functions? These functions have two inverse operations. If we have a function

$$c = b^a$$

and we know a and c we can ask, What is b ? It is the a th root of c , or

$$b = \sqrt[a]{c}$$

On the other hand, we might suppose we know b and c and ask, What is a ? It is defined to be the logarithm in base b of c , or

$$a = \log_b c$$

assuming b and c to both be positive.

But we run into another problem because the set of rational numbers is not closed under taking roots. For example, it is impossible to solve this equation within the set of rational numbers:

$$x^2 = 2$$

That is, it is impossible to find two integers whose division yields $x = \sqrt{2}$. The proof of irrationality of $\sqrt{2}$ is found “in Euclid,” but is traditionally attributed to Hippasus, a member of the Pythagoreans, who discovered it while the Pythagoreans were sailing to Sicily to escape political unrest in Crete. The Pythagoreans made a religion of the worship of how rational numbers could explain the harmonies of music and planetary periods of planets (or so they thought). Consequently, they were horrified that such numbers could exist. At first they tried to ignore them, but it is patently obvious that $\sqrt{2}$ could be constructed as the hypotenuse of a right triangle with sides equal to 1 unit using, ironically, the Pythagorean Theorem! The hypotenuse clearly has a precise, finite length, but its length cannot be measured within the set of rational numbers. Legend has it that Hippasus was promptly thrown overboard and the Pythagoreans swore themselves to secrecy about irrational numbers; the penalty for transgression was death.

This proof is one of the most beautiful in all of mathematics and relies on a technique known as “proof by contradiction.” That is, we begin by assuming that something is true and then proceed logically to show that this leads to a contradiction. Therefore, our original assumption that the statement was true was in fact false and

so its converse must be true. “Proof by contradiction” is in some ways similar in spirit to beginning an experiment by accepting the null hypothesis, then showing empirically that it leads to contradictions with observation so we then reject it and accept the alternative hypothesis. Here is the proof:

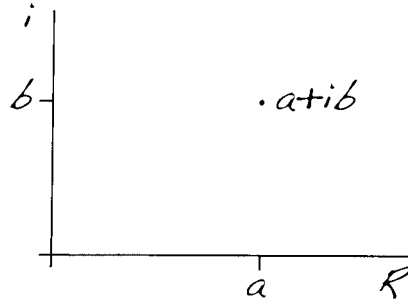
Suppose $\sqrt{2} = m/n$, where m and n are natural numbers that have no common divisor; m/n is therefore rational. It follows that $2n^2 = m^2$ which means that m^2 is even because it is twice another number and m is therefore even because the square of an even number is an even number (an even number contains 2 as a divisor and squaring 2 equals 4, another even number which is a divisor of the squared result). Then $m = 2p$, $m^2 = 4p^2$. Therefore, $2n^2 = 4p^2$, or $n^2 = 2p^2$, which means that n is even. But if n and m are both even, this contradicts our original assumption that $\sqrt{2}$ is rational and can be expressed as the quotient of two numbers with no common divisor – we have shown that they must both have 2 as a common divisor. Therefore $\sqrt{2}$ is not rational. Q.E.D. (*Quod erat demonstrandum* – “that which was to be demonstrated”).

Numbers that are not rational are called irrational (which literally means “cannot be expressed as a ratio”). The decimal expansion of an irrational number has neither a repeatable pattern nor does it end in 000 . . . The rationals and irrationals together form the real numbers. While real numbers are needed for closure under the operation of taking roots, we are now beyond the numbers we need simply to collect data. Data can only be expressed as rational numbers because it is impossible to write down the infinite expansion of an irrational number in a field or laboratory notebook (or in a spreadsheet, for that matter), despite the demonstrable construction of real lines whose lengths are irrational numbers, such as the length of a hypotenuse of a right triangle with sides = 1.

Let us now make three arbitrary rules in order to avoid ambiguities when we multiply numbers of different sign (Courant and Robbins 1961). The first rule is that multiplication by two positive real integers always yields a positive integer, that multiplication by two negative integers always yields a positive integer, and that multiplying a positive real by a negative real always yields a negative integer. We can extend this to multiplication of any two real numbers simply by factoring out a 1 or -1 , multiplying the two (now positive) real numbers together, applying these rules to the 1 and/or -1 which have been factored out, and finally multiplying the second resulting product (1 or -1) by the product of the two positive real numbers. Thus, the root of any real number has both positive and negative solutions. Normally we are not interested in the negative answer since there are no negative organisms, but there will be times when these two solutions to an equation will be useful.

Here we enter our next-to-last closure problem. What is $\sqrt{-1}$? By the above rules, there are no two numbers alike that would yield -1 as their product. This means that we cannot solve equations such as $x^2 + 1 = 0$. We therefore define a new number to be equal to $\sqrt{-1}$ and call it i . Real numbers that are multiplied by i are called imaginary numbers, and numbers that are the sum of real plus imaginary numbers are called complex numbers of the form $a + ib$ where a and b are real numbers. The set of real numbers is a subset of complex numbers where $b = 0$. Complex numbers reside, not along a line as in the real numbers, but in the complex plane (Fig. 2-1) composed of an axis for the real part (R) and a perpendicular (orthogonal) axis for the imaginary part (i).

Fig. 2-1 The complex plane.



Unlike real numbers, which lie to the right or left of each other along a line and so we can ask which is the larger of two numbers (meaning, which lies to the right of the other), complex numbers cannot be ordered – it makes no sense to ask which of two complex numbers is the larger. Fortunately, we do not need complex numbers to measure quantities, but we do need them to get solutions to some algebraic equations. We shall see how they play an important role in the emergence of periodic solutions to equations which bear strong analogies to population oscillations.

Two complex numbers that differ only in the sign of the imaginary part are called complex conjugates, for example $a + ib$ and $a - ib$. Complex conjugates are equally valid solutions to equations so long as we change positive imaginary parts to negative imaginary parts everywhere else in the equation.

How do we perform the arithmetic operations on complex numbers? To add or subtract two complex numbers simply add the corresponding real and imaginary parts. Similarly, to multiply any two complex numbers simply multiply through the parentheses just as you would do to multiply $(a + b)(c + d)$ for any four real numbers, except that now when you multiply the two terms containing i you are squaring i and so you get a real number times -1 for that term. Here is a demonstration which is also a proof for closure of complex numbers under multiplication:

$$(r + is)(p + iq) = rp + irq + isp + i^2sq = (rp - sq) + i(rq + sp)$$

which is another complex number because its components $(rp - sq)$ and $(rq + sp)$ are reals. To divide two complex numbers, you multiply both numerator and denominator by the complex conjugate of the denominator, yielding another complex number:

$$\frac{a + ib}{c + id} = \frac{a + ib}{c + id} \frac{c - id}{c - id} = \left(\frac{ac + bd}{c^2 + d^2} \right) + i \left(\frac{bc - ad}{c^2 + d^2} \right)$$

The set of complex numbers is closed under all algebraic operations (addition, multiplication, their inverses subtraction and non-zero division, powers, roots, and logarithms). We now finally have a set of mathematical objects that can stand for the magnitude of any biological object (e.g., population density) but which are closed under all operations that will stand for biological processes (e.g., reproduction, death, etc.). Although we think of the magnitudes of biological processes (rates, for example) in terms of the real numbers and we measure them using rational numbers, it is in the domain of complex numbers that we can perform all mathematical operations

that stand for biological processes. Most specifically, the complex domain allows us to solve all polynomial equations in what is known as the fundamental theorem of algebra (but which, as noted by Courant and Robbins (1961), should be called the fundamental theorem of complex numbers): Every polynomial equation of degree n :

$$f(x) = x^n + A_{n-1}x^{n-1} + \cdots + A_1x + A_0 \quad (2.1a)$$

can be factored into the product of exactly n factors:

$$f(x) = (x - a_1)(x - a_2) \cdots (x - a_n) + a_0 \quad (2.1b)$$

where a_n are complex numbers that are solutions to the equation $f(x) = 0$. Many of the equations we shall encounter involve polynomials of the form of Eq. 2.1a.

The absolute value of a complex number, called the modulus and symbolized as $|x + iy|$, is the distance from the origin to the point representing that number in the complex plane. By the Pythagorean theorem, this is:

$$|x + iy| = \sqrt{x^2 + y^2} \quad (2.2)$$

The product of two complex conjugates is the square of their modulus:

$$(x + iy)(x - iy) = x^2 - ixy + ixy - i^2y^2 = x^2 + y^2 \quad (2.3)$$

The set of all numbers $x + iy$ for which $(x + iy)(x - iy) = r$ define a circle in the complex plane of radius r . Therefore, the product of complex conjugates (or the square of their modulus) defines a circle in the complex plane.

Now, how do x and y vary as one moves around a circle in the complex plane with radius arbitrarily set to 1 unit (i.e., as θ varies from 0 to 2π , where θ is the angle the radius makes with the x -axis)? x varies as the cosine and y varies as the sine of θ . Therefore, $x = \cos(\theta)$ and $y = \sin(\theta)$. We can now state a formula sometimes attributed to Leonhard Euler (1707–1783)* to express Eq. 2.3 as a complex power of e :

$$e^{i\theta} = \cos \theta + i \sin \theta \quad (2.4)$$

Since the cosine and sine functions oscillate, so does $e^{i\theta}$. In addition, $e^{a+i\theta} = e^a e^{i\theta} = e^a(\cos \theta + i \sin \theta)$. Now, θ need not necessarily be an angle but could stand for any real number. So solutions of equations that are of the form $e^{a+i\theta}$ will exhibit oscillations for any variable which is measured by real numbers. If these solutions represent population densities, then the model predicts the conditions under which populations will also oscillate. As we shall see, complex numbers will become essential to understanding the solutions of differential equations which are analogs to predator-prey cycles.

*There is some recent evidence that this formula was known before Euler. Roger Cotes (1682–1716), the editor and the author of the preface of the second edition of Newton's *Principia*, apparently wrote the logarithmic transformation of Eq. 2.4 in a notebook in 1714 when he would have been 32 and Euler 7 years of age (Friedlander 2007).

Through successive steps of developing a system of numbers that is closed under all operations we wish to use, we end up with the system of complex numbers which solves all our problems of closure under all arithmetical operations analogous to biological processes except for one, the problem of dividing by intervals that are so infinitesimally small that they approach zero, for which division is undefined. To deal with this problem requires the development of a precise concept of limits and the calculus, to which we now turn.

Limits, calculus, Taylor series

Much of what we deal with in ecology are rates of change of biological objects: growth of an organism, decay of a dead leaf, fluctuations in populations, accumulation or erosion of soil, increases or decreases in lake levels, etc. But rates of change are some of the hardest things to measure. What we measure are static properties such as the sizes of objects at different times and then infer that change has taken place between those two measurements. But what exactly is meant by “change” here? How do we logically go from two static measurements to a dynamic quantity representing “change”? How can we express mathematically the rate of change of a quantity and, knowing this, how do we find the quantity at any time? Does the change happen continuously or in discrete jumps? Does it make a difference? How do continuous change and discrete change relate to one another? We need to think clearly about rates of change of populations and ecosystems in order to better understand what our measurements mean.

When we calculated the rate of growth of the bee colony by dividing b new individuals counted between t_0 and t_1 by dividing by $t_1 - t_0$, we made an estimate of the average rate of change over a finite time interval. What if we wanted to know the rate of population growth not over some time interval but exactly at a specific point in time? That is, what is the instantaneous rate of change of a population? This is not something that can be measured because measurements have to take place over a finite length of time, during which instantaneous rates may be changing. How do we deal with this problem?

The mathematical theory that is the analog of this biological problem is calculus, specifically differential equations. Although Isaac Newton (1643–1727) and Gottfried Leibniz (1646–1716) provided the first attempt at a rigorous formalization of the calculus, their work was the culmination of many centuries of grappling with the problem of the limit of a series of repeated operations or the limit of a function that contains an infinite series of terms (Boyer 1959). The limit of a sequence or function is the fundamental concept underlying calculus. This is a very subtle concept which required considerable research over several centuries to refine precisely. But since most of the equations of mathematical ecology are framed as differential equations and since a very important property of these equations, their stability, draws on the idea of a limit, we need to grapple with the nature and problems of this concept.

Suppose

$$f(x) = ax^n$$

What is the rate of change of $f(x)$ at x ? Add a small increment Δx to x and find the new value of $f(x + \Delta x)$:

$$f(x + \Delta x) = a(x + \Delta x)^n$$

Suppose $n = 2$:

$$f(x + \Delta x) = a(x^2 + 2x\Delta x + \Delta x^2) = ax^2 + 2ax\Delta x + a\Delta x^2$$

The net change in $f(x)$ can be found as follows:

$$f(x + \Delta x) - f(x) = \Delta f(x) = 2ax\Delta x + a\Delta x^2 + ax^2 - ax^2 = 2ax\Delta x + a\Delta x^2$$

The average rate of change in $f(x)$ over the interval Δx is

$$\frac{\Delta f(x)}{\Delta x} = \frac{2ax\Delta x + a\Delta x^2}{\Delta x} = 2ax + a\Delta x$$

What happens as $\Delta x \rightarrow 0$? The term $a\Delta x \rightarrow 0$ and so the right hand side $\rightarrow 2ax$. (From this point on, we will abbreviate “right hand side” as r.h.s.) At this point, Leibniz (and Newton, too, but using other notation) renamed $\Delta f(x)/\Delta x$ as $df(x)/dx$ and call it a derivative of f with respect to x .

But there is a problem here, first pointed out by George Berkeley (1685–1753), an Anglican Bishop and philosopher and contemporary of Newton. Berkeley’s criticism is not to be taken lightly and in fact addressing it took several centuries of mathematical research during which much new mathematics was created. Berkeley’s criticism (a logically correct one) was the right hand side (r.h.s.) of the above equation may go to $2ax$ as $a\Delta x$ goes to zero, but it appears that the left hand side (l.h.s.) is undefined because there we divide by zero. This is a contradiction, so the calculus is not logically consistent (closed in the real numbers on the right but not on the left). Berkeley called $a\Delta x$ the “ghost of a departed quantity.”

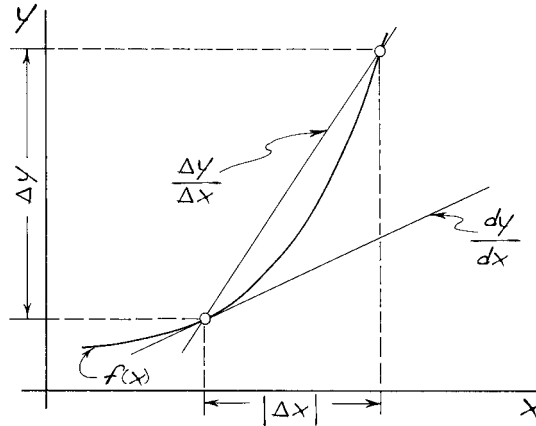
However, calculus worked so well in precisely and accurately describing the dynamics of much of the physical world, from the fall of apples, to vibrations of violin strings, to the tides, to the entire solar system, that mathematicians and physicists did not at first worry too much about this problem. But eventually mathematicians could not avoid Berkeley’s criticism. What if Berkeley was correct? Then eventually all science based on calculus would run into a crisis when it tried to predict some phenomenon and failed because of this flaw. The concept of the “limit” was developed to avoid this crisis in the foundations of calculus.

Around 1820, Augustin-Louis Cauchy (1789–1857) provided a logical definition of the limit that avoids Berkeley’s criticism altogether by precisely defining what it means to be “arbitrarily close” to a number. The concept of “arbitrarily close” (which we have also encountered in the proof of the infinite density of rational numbers) is the key to the concept of limits, which is the fundamental concept of calculus and analysis. Cauchy defined a limit as:

$$\begin{aligned} &\text{For any } \varepsilon > 0, \text{ there is a corresponding } \delta > 0 \text{ such that if} \\ &0 < |\Delta x| < \delta, \text{ then } |f(x) - L| < \varepsilon \end{aligned} \tag{2.5}$$

That is, if you give me an ε , no matter how small, which is greater than the difference between a function and its limit, I can produce another small number δ which

Fig. 2-2 The slope of a function, $y = f(x)$, is a limit. For every $|\Delta x| < \delta$, one can find an ε such that $|\Delta y/\Delta x - dy/dx| < \varepsilon$.



depends on ε and within which I can get a Δx that makes the magnitude of the difference between $f(x)$ and L smaller than ε . If you then give me a smaller ε , I can again produce an even smaller interval of size δ within which I can get a smaller Δx that makes the magnitude of the difference between $f(x)$ and L even smaller. The intervals that are of magnitude ε and δ bracket $|f(x) - L|$ and Δx , respectively, in the same way that the rational numbers q and p bracketed another rational number q' as we saw above. In this way, I can get $f(x)$ arbitrarily close to L just as we got q' arbitrarily close to q by constructing smaller and smaller intervals around q that still contained q' . Now consider the instantaneous slope (dy/dx) of a function $y = f(x)$. Again, we can get arbitrarily close to this by finding a $\Delta x < \delta$ such that $|\Delta y/\Delta x - dy/dx| < \varepsilon$, where ε is as small as you like. The limit is what $\Delta y/\Delta x$ gets closer to as Δx gets smaller and smaller (Fig. 2-2).

Most people feel something like Alice falling through the looking glass when they encounter this definition. Part of the reason is that whereas before we started with the notion that the independent variable x was approaching something, then asked what happened to the dependent variable $f(x)$, we now start with some conditions on $f(x)$, namely that it is within some difference ε of L at a given x then ask how much do we have to change x to make the difference between $f(x)$ and L even smaller? The answer is $\Delta x < \delta$. Cauchy's trick was to reverse the order in which we consider the so-called dependent ($f(x)$) and independent (x) variables from the order that Newton and Leibniz considered them.

This is highly abstract notion, but the advantage is that it avoids ever getting into the situation of dividing by 0 by giving an explicit "static" procedure for finding the limit of a function and avoiding the metaphysical notion of "motion" of a variable towards something else. The intuitive idea of motion is a physical one which cannot be mathematically defined without running into difficulties such as dividing by 0 when the variable "arrives" at a "place" in any physical sense (Courant and Robbins 1961). Here, the trick is not to think of ∞ or 0 as numbers that x or Δx eventually reach, but rather think of the symbols $\rightarrow \infty$ or $\rightarrow 0$ as representing a never-ending process that we can terminate when $f(x)$ is within some arbitrarily small difference ε of a

limit. There is no reason to give up the intuitive feeling implied by $x \rightarrow \infty$ or $\Delta x \rightarrow 0$ as we begin to translate biological ideas into mathematics, but eventually we must make sure that we are rigorous by translating these intuitive ideas into Cauchy's definition.

The fact that most people encounter this definition of limits in elementary calculus courses in high school or in their freshman year in college should in no way indicate that it is trivial. Rather, the reason why it is encountered so early in mathematics curriculum is because it is foundational to almost all of higher mathematics. It took mathematicians almost two centuries to rid the concept of limits from the undefined quantities that Berkeley said plagued them, and even today the concept of limit still proves enormously fruitful (Hight 1977). And, as we shall see, the concept of limit leads to a very rich and fruitful understanding of the behavior of ecological systems near their equilibria, one of the core concepts of mathematical ecology.

Incidentally, now that a limit is precisely defined, we can now see where Berkeley went wrong in his criticism. Berkeley assumed that

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta f(x)}{\Delta x} = \frac{\lim_{\Delta x \rightarrow 0} \Delta f(x)}{\lim_{\Delta x \rightarrow 0} \Delta x}$$

which is false.

We can also use this definition of a limit to clarify what is meant by the concept of continuity. We have been assuming that the domain of differentiable functions is the set of real numbers. A function f is continuous at a point x if $|f(x) - f(p)| < \epsilon$ whenever $|x - p| < \delta$. Because the set of real numbers is closed under the operations used in these functions, the continuity (or density) of real numbers is in some sense equivalent to the continuity of a function for which this limit holds. Most of the functions we will be using are not only continuous, but also continuously differentiable – that is, they have a unique and well-defined slope at every point given by the derivative evaluated at that point. Functions with corners are continuous but not continuously differentiable because the slope at the corner is not defined. Further discussion of this deep relationship between real numbers, continuous functions, and limits can be found in Courant and John (1965). Table 2.1 is a list of useful rules for finding derivatives of general forms of functions that we will frequently encounter. These have been all derived rigorously using Cauchy's definition of the derivative as a limit of $\Delta y/\Delta x$ as $\Delta x \rightarrow 0$ (see any calculus text).

A particularly important function in Table 2.1 to note are the exponential functions, $y = e^x$. An exponential function is one whose slope at any point x equals the value of the function at x . Thus, if the function increases (or decreases) with x , the rate of change also increases (or decreases) proportionally. Therefore, the rate of change of the exponential function continues to increase faster and faster (or decrease slower and slower) with increases (decreases) in x .

A set of equations that we will often use are called differential equations. These are equations that specify the rate of change (dx_i/dt) of some variable, x_i , with time, t , as functions of t , x_i , perhaps some other variables, $x_2, x_3, \dots$, and a set of parameters (constants: $a, b, c, \dots$) which act on the variables by means of arithmetic operations: $dx/dt = f(t, x_1, x_2, x_3, \dots; a, b, c, \dots)$. Although t appears as an argument

Table 2.1 Useful rules of differentiation.

Function: $y = f(x)$	Derivative: dy/dx
<i>Summation Rule</i> $y = f(x) + g(x)$	$dy/dx = df(x)/dx + dg(x)/dx$
<i>Product Rule</i> $y = f(x)g(x)$	$dy/dx = f(x) \cdot dg(x)/dx + g(x) \cdot df(x)/dx$
<i>Quotient Rule</i> $y = f(x)/g(x)$	$dy/dx = [df(x)/dx \cdot g(x) - dg(x)/dx \cdot f(x)]/[g(x)]^2$
<i>Power Rule</i> $y = cx^n$	$dy/dx = cnx^{n-1}$
<i>Chain Rule</i> $y = g(f(x))$	$dy/dx = dg/df \cdot df/dx = dg/dx$
<i>Constants</i> $y = c$	$dy/dx = 0$
<i>Linear functions</i> $y = cx$	$dy/dx = c$
<i>Trigonometric functions</i> $y = \sin x$ $y = \cos x$	$dy/dx = \cos x$ $dy/dx = -\sin x$
<i>Exponential functions and the Natural Logarithm</i> $y = a^x$ $y = \ln x$ $y = e^x$ $y = e^{u(x)}$	$dy/dx = (\ln a) a^x, a > 0$ $dy/dx = 1/x, x > 0$ $dy/dx = e^x$ $dy/dx = e^{u(x)} du(x)/dx$

in this general form, it need not always appear in the list of arguments. Differential equations in which t is not an argument are called autonomous differential equations. As we shall see, it is often easiest to begin constructing a model by writing the autonomous differential equation(s) for the rates of change. This is often the easiest step in modeling – analyzing those equations is the hard part!

When f depends not just on a single variable, x , but on several variables $x_1, x_2, \dots$, we can calculate the change in f for an infinitesimal change in each variable while holding all others constant by taking the partial derivative with respect to that variable alone. For example, if

$$f(x_1, x_2) = x_1^a x_2^b$$

then

$$\frac{\partial f}{\partial x_1} = ax_1^{a-1} x_2^b$$

$$\frac{\partial f}{\partial x_2} = x_1^a bx_2^{b-1}$$

A partial derivative is analogous to a partial correlation coefficient or the main effect in an ANOVA. When we analyze models of interactions between two or more species to find out how each species affects not only itself but also the others we will have to take partial derivatives of their differential equations with respect to each species.

The gradient of a function of several variables is a useful feature of partial derivatives. The gradient of a function is a vector of the partial derivatives of that function for each variable.

$$\nabla f(x_1, x_2, x_3) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right] \quad (2.6)$$

If the gradient of a function is evaluated at some point (x_1, x_2, x_3) , then the gradient always points in the direction of the greatest increase of that function.

So now you know how to find the rates of change of a function with respect to one or several variables. What about the opposite problem: if we write down an equation for dx/dt , how do we find x at some time t ? This is the inverse of differentiation and is called integration. Differentiation gives the slope at a point, or $dy/dx = f(x)$; integration gives the cumulative change in y over x , which is the same as the area under the curve dy/dx from x_1 to x_2 . The Fundamental Theorem of Calculus relates these two inverse processes by giving the limit of the sum of a series of n polygons in the interval $[a, b]$ each of width Δx and heights of $f(x_i)$ and $f(x_i + \Delta x)$:

$$\lim_{n \rightarrow \infty, \Delta x \rightarrow 0} \sum_{i=1}^n f(x_i) \Delta x = \int_a^b f(x) dx = F(b) - F(a) \quad (2.7)$$

where $\Delta x = (b - a)/n$, x_i is any sample point in the i^{th} subinterval of $[a, b]$, and $dF(x)/dx = f(x)$.

Integration is the opposite of differentiation. For example to find the derivative of a power function, do this:

$$f(x) = x^a$$

$$\frac{df(x)}{dx} = ax^{a-1}$$

but to find the integral of the same function do this:

$$\int x^a dx = \frac{1}{a+1} x^{a+1} + C$$

for $a \neq -1$. When $a = -1$,

$$\int x^{-1} dx = \int \frac{1}{x} dx = \ln|x| + C$$

where C is known as the constant of integration. When we integrate a function, we must add a constant to the solution because there are an infinite number of functions that have the same slope at all points but differ only by a constant. Specifying initial conditions (values of x at $t = 0$) when integrating or solving differential equations specifies the constant.

If we know the initial conditions at some time t_0 , which must be specified by measurement or by arbitrarily assigning some value to x at t_0 , then we can approximate its behavior by using the derivative to project the time course of x forward from t_0 to t :

$$x(t) = x(t_0) + (t - t_0) \left. \frac{dx}{dt} \right|_{t_0}$$

But unless the derivative is constant for all t , this will only work in the immediate neighborhood of t_0 . If the derivative itself changes, then you must take into account the rate of change of the derivative, or the second derivative, and add a term incorporating this:

$$x(t) = x(t_0) + (t - t_0) \left. \frac{dx}{dt} \right|_{t_0} + \frac{(t - t_0)^2}{2!} \left. \frac{d^2x}{dt^2} \right|_{t_0}$$

and we can continue adding terms in higher orders of derivatives as needed for the desired accuracy:

$$x(t) = x(t_0) + (t - t_0) \left. \frac{dx}{dt} \right|_{t_0} + \frac{(t - t_0)^2}{2!} \left. \frac{d^2x}{dt^2} \right|_{t_0} + \frac{(t - t_0)^3}{3!} \left. \frac{d^3x}{dt^3} \right|_{t_0} + \dots \quad (2.8)$$

This series is called a Taylor Series or a Taylor expansion and was discovered by Brook Taylor (1685–1731), a contemporary of Newton. This series is very extensively used in any applications of mathematics to any natural phenomenon, including in theoretical ecology.

If you have good empirical estimates of the values of the parameters upon which the function depends and the initial conditions, then numerical simulations can be made using the initial value and the second term of the Taylor Series. Define $t - t_0$ be a constant step of interval h . Then the next point on the curve, $(x(t + h), t + h)$ can be approximated from any point $(x(t), t)$ starting from the initial conditions $(x(t_0), t_0)$ using only the first derivative:

$$x(t + h) = x(t) + h \left. \frac{dx}{dt} \right|_{t=t_0} \quad (2.9)$$

This is Euler's method of order 1 (i.e., uses only first derivative). The error of Euler's method can be large, especially if the step size is large and/or if the second or higher order derivatives evaluated at $x(t)$ are large (Euler's method ignores the higher order derivatives). The error is reduced by making the step size very small, or by incorporating more information (more terms of Taylor Series expansion that include the higher order derivatives).

Another method is the Runge–Kutta method, after Carle Runge (1856–1927) and Martin Kutta (1867–1944), two collaborators who developed it. This corrects for the error in Euler’s method by calculating the derivatives at $x(t)$ and the one at $x(t + h)$ then averaging these two derivatives to project the point forward along the function. In other words:

$$x(t + h) = x(t_0) + h \frac{\left(\frac{dx}{dt} \Big|_{t_0} + \frac{dx}{dt} \Big|_{t_0+h} \right)}{2} \quad (2.10)$$

This is called Runge–Kutta Order 2. Runge–Kutta Order 4 uses the slopes at t , $t + h$, and $t + h/2$ to correct for subsequent changes to the value of dx/dt near t . This is even more accurate and is the standard default method for most programs that find numerical approximations of solutions. A rule of thumb for the step size for Euler’s method (the worst case, and therefore the largest step size) is that the step size times the derivative should not exceed 0.5 over the range of data or interval of time of interest. Because they use more information about the function than Euler’s method, Runge–Kutta methods can use larger step sizes.

Exercise 2.1

Write out the first four terms of the Taylor Series expansion of:

$$x(t) = x(0)e^{kt}$$

Hint: recall from Table 2.1 that $d(e^u)/dt = e^u du/dt$ and that here $u = kt$.

Given that $x(0) = 1.0$ at $t(0) = 0$ and $k = 0.5$, evaluate the Taylor Series expansion for $x(t)$ to four terms for $t = 1$. How close is this to simply finding $x(t)$ by plugging in $t = 1$ and $x(0) = 1$ on a calculator and using the e^x key?

Matrix algebra and linear systems

In ecology, we are often interested in exploring the behavior of whole systems of species or ecosystem composed of individual components which interact through biological processes. We are interested not simply in the dynamics of each species or component in isolation, but the dynamics of each species or component in the context of all the others and how those coupled dynamics account for properties of the system as a whole, such as its persistence. This is what people seem to mean when they say that ecology is “holistic,” an otherwise rather vague term.

To address this problem mathematically, we will construct systems of coupled equations, one equation for each species or ecosystem component whose rate of change is a function of both its size and also the size of at least one other species or component. We are interested not simply in the behavior of each equation representing the dynamics of individual species or components, but the behavior of the collection of all equations operating simultaneously, just as we are interested in the behavior of a system of species operating together in the real world.

In order to understand properties of whole systems and how they depend on the properties and interactions of each part, we need a mathematical object which treats a collection of mathematical objects simultaneously. A matrix fills this need. A matrix

is a mathematical object in which numbers, functions, or other mathematical objects are arranged in rows and columns. Matrices can model how the components of a system interact with each other, thereby allowing one to ask certain questions about the collection of objects as a whole, especially if the objects interact according to specified rules. Inferences about the stability of an ecological system can be made from some well-defined properties of matrices.

The modern theory of matrix algebra was formalized by Arthur Cayley (1821–1895), William Rowan Hamilton (1805–1865), and James Joseph Sylvester (1814–1897). However, the mathematical idea of a matrix goes back to 250 BC in the publication of the Chinese text *Nine Chapters on the Mathematical Art* by an anonymous mathematician in which bamboo sticks representing equations are manipulated on a checkerboard, the checkerboard being a physical representation of a matrix with its squares standing for a particular address uniquely specified by its row and column.

A matrix is characterized by the number of its rows $\times$ the number of its columns, just like the dimensions of lumber are characterized by width and thickness (e.g., a 2×4 in American lumber standards). The standard notation is for m rows and n columns. Here are some examples:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad [1]$$

$2 \times 2 \qquad 2 \times 3 \qquad 3 \times 2 \qquad 3 \times 1 \qquad 1 \times 1$

Note that, even though they contain the same elements, the second and third matrices are not the same because the elements are in different positions. A matrix is often symbolized by a bold capital letter, e.g., **A**. The position of an object in a matrix is given by its row and column numbers, usually subscripted i, j . The next-to-last matrix above is simply a single column of objects and is called a vector. Vectors are often symbolized by bold small letters, e.g., **v**. The last matrix contains a single object and is called a scalar. Scalars are symbolized by small, unbold letters, often Greek letters, e.g., λ .

To add or subtract two matrices, simply add or subtract the corresponding elements:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

It is obvious that doing this requires that the two matrices contain the same number of rows and columns in order to achieve a complete 1:1 pairing of each element in one matrix with a corresponding element in another. The identity element for addition/subtraction of a matrix is one of the same number of rows and columns but in which each element is 0.

Multiplying matrices follows the rule “row times column, then add them up.” That is, take each element in row i of the first matrix, multiply it by the corresponding elements in column j of the second matrix, add these products up, and place the result in position ij in the new product matrix:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i}b_{i1} & \sum_{i=1}^n a_{1i}b_{i2} & \cdots & \sum_{i=1}^n a_{1i}b_{ip} \\ \sum_{i=1}^n a_{2i}b_{i1} & \sum_{i=1}^n a_{2i}b_{i2} & \cdots & \sum_{i=1}^n a_{2i}b_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{mi}b_{i1} & \sum_{i=1}^n a_{mi}b_{i2} & \cdots & \sum_{i=1}^n a_{mi}b_{ip} \end{bmatrix}$$

You can multiply two matrices only if the number of columns in the first equals the number of rows in the second. In general, the commutative law does not apply to matrices.

The product of a square $n \times n$ matrix and a vector of n rows is a compact way of depicting a set of simultaneous linear equations:

$$\begin{aligned} a_{11}x_{1,t} + a_{12}x_{2,t} + \cdots + a_{1n}x_{n,t} &= x_{1,t+1} \\ a_{21}x_{1,t} + a_{22}x_{2,t} + \cdots + a_{2n}x_{n,t} &= x_{2,t+1} \\ &\vdots \\ a_{m1}x_{1,t} + a_{m2}x_{2,t} + \cdots + a_{mn}x_{n,t} &= x_{m,t+1} \end{aligned}$$

in more compact matrix form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_t = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}_{t+1}$$

The change in the vector $\mathbf{x}_t$ to $\mathbf{x}_{t+1}$ depends entirely on the elements of the matrix, a_{ij} . We will see that we have to analyze the properties of this matrix to determine the dynamics and stability of many models.

The product of a matrix and a vector is the same as the sum of the products of each column in the matrix times the corresponding row vector element:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n$$

The identity element for multiplication has a main diagonal filled with 1's and all other elements equal to zero:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

This is a particularly useful matrix often symbolized by $\mathbf{I}$.

In this book, we will work exclusively with “square” matrices where the number of rows equals the number of columns. Such $n \times n$ matrices will articulate the interactions

of each of n elements (age classes, species, land cover classes, ecosystem components, etc.) with itself and with each of the other elements. Square matrices have several properties that will help us understand properties of the whole biological system of which the matrix is a mathematical analog. The simplest property is the trace, which is the sum of all diagonal elements, symbolized by trA or sometimes by τ :

$$trA = \sum_{i=1}^n a_{ii} \quad (2.11)$$

The determinant of a matrix A is a scalar (pure number) that is also a useful property. Determinants were first discussed in 1683 by Takahazu Seki Kowa (1642–1708), a Japanese mathematician born into a samurai family, in his treatise *Fukudai no Ho*. The determinant of a matrix is symbolized by a Δ , by $\det A$, or by $|A|$. The determinant of a 2×2 matrix is calculated as:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (2.12)$$

The determinants of larger matrices are found by a method called Laplace expansion, or expansion by minors. The minor of an element a_{ij} is a matrix formed by crossing out the row i and column j . For a 3×3 matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

the minor of a_{11} is what remains after crossing out row 1 and column 1:

$$\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

The determinant of a 3×3 matrix is calculated by choosing any row and multiplying each of its elements by the determinants of their minors, and subtracting and adding each result in alternate succession. For example, if we choose row 1, then

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Can you see that the more succinct formula for the determinant of a 2×2 matrix (Eq. 2.12) is exactly this formula? It is best to use computer programs to calculate determinants of larger matrices, although this algorithm still applies. You just have to apply it in sequence until you get to minors that are determinants of 2×2 matrices and then apply the above formula for the determinant of a 2×2 matrix.

The value of a determinant is unaffected if all rows and columns of a matrix are transposed (rows made into columns and columns made into rows). However, the value of a determinant retains its numerical value but reverses its sign if any two rows or columns are interchanged.

The multiplicative inverse of a matrix is another matrix which, when multiplied by the first matrix, yields the identity matrix:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

The inverse of a matrix exists only if $\det \mathbf{A} \neq 0$. The inverse of a 2×2 matrix is:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

More general formulas exist for larger square matrices, but it is best to use a computer program to compute them.

An inverse does not exist if $\det \mathbf{A} = 0$. For example, in the above formula for a 2×2 matrix this would lead to division by zero, or a singularity. Matrices without an inverse are therefore called singular. If you use a computer program to calculate the inverse of such a matrix, you will often get back an error message something like: "Singularity encountered." If a matrix cannot be inverted, then the set of simultaneous equations it represents may not have a solution. If the elements of a matrix hypothesized to articulate the interactions among species has a determinant equal to 0, then this implies that either the model is not a good description of the ecological system (most likely), or else evolution has selected against systems of these species (less likely, but a more interesting result).

If you know the inverse of a matrix, you can solve simultaneous equations. For example, if

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

and $\mathbf{b}$ is a vector of known (measured) constants, then we can solve for $\mathbf{x}$ by:

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

$$\mathbf{I}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Matrix operations are linear operations on vectors. A function is an operation which is linear in a and b if both of the following can be shown:

$$f(a + b) = f(a) + f(b) \quad (2.13a)$$

and

$$cf(a) = f(ca) \quad (2.13b)$$

where c is a constant. The first condition says the behavior of a system $f(a + b)$ equals the sum of the behavior of each of its parts $f(a) + f(b)$. In other words, "the system is the sum of its parts." The second condition says that a change of magnitude c in the size of a part of a system, a , yields a proportional change in the behavior of the system. In other words, proportional changes to the size of a system's parts yield proportional changes to its behavior.

More familiarly, linear algebraic operations can be transformed into

$$y = ax + 0$$

which is linear in x because if we define x as the sum of two “parts” (e.g., $x_1 + x_2$) then

$$y = a(x_1 + x_2) = ax_1 + ax_2$$

and so condition (2.13a) is satisfied. Similarly,

$$y = c(ax) = a(cx)$$

and so condition (2.13b) is satisfied.

In the above, c is a scalar and x and y could either be scalars or they could be vectors $\mathbf{x}$ and $\mathbf{y}$. If they are vectors, then a is a matrix $\mathbf{A}$. Therefore, matrix operations are linear because

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$$

and

$$c\mathbf{A}\mathbf{x} = \mathbf{A}c\mathbf{x}$$

If the “ y -intercept” does not equal 0, then we can always translate y to y' :

$$y = ax + b$$

$$y' = y - b = ax$$

Equations in which x is raised to a power, such as

$$y = ax + bx^2$$

are not linear in x because a proportional change in x results in a change in y that is greater than that proportional change:

$$a(cx) + b(cx)^2 > c(ax + bx^2)$$

For $c, b > 0$. In addition, if we subdivide x into two parts, x_1 and x_2 , then the behavior of the sum of the parts is greater than the sum of the separate behaviors of the parts:

$$a(x_1 + x_2) + b(x_1 + x_2)^2 > (ax_1 + bx_1^2) + (ax_2 + bx_2^2)$$

(square the term $(x_1 + x_2)$ on the l.h.s., then multiply through the terms in parentheses by a and b and you will see that the l.h.s differs from the r.h.s. by bx_1x_2).

Functions that involve powers of a single variable or products of two or more variables violate the requirements of a linear system because neither condition is

satisfied. Such equations (as well as others that do not meet these conditions) are called nonlinear, although sometimes ax is called the linear term and bx^2 is called the nonlinear term. Therefore, in nonlinear systems proportional changes in the size of parts can produce disproportional changes in the behavior of the system and the behavior of the system is “more than the sum of its parts” (as is often and vaguely stated for ecosystems).

We shall see that linear and nonlinear equations have very different properties, but also that under certain circumstances linear equations can be very useful approximations to nonlinear equations, which are always much more difficult to solve. In particular, the first two terms of the Taylor Series expansion of a function:

$$x(t) = x(t_0) + (t - t_0) \left. \frac{dx}{dt} \right|_{t_0}$$

can be put into linear form if we translate $x(t)$ by subtracting $x(0)$. The derivative then is equivalent to a in the above definition. This makes sense because a normally is interpreted as standing for the slope of x with respect to t , which is the definition of the derivative of x with respect to t .

Statistics is the analysis of linear hypotheses which are tested by experiments designed to meet the assumptions of the statistical method used. Most people are familiar with the assumption of normality, but an even more fundamental assumption of statistics is that the system is linear. All of statistics is based on the expectation of a random variable $E(X)$, or

$$E(X) = \sum_{i=1}^n x_i f(x_i)$$

where x_i is a value and $f(x_i) = [\text{Pr}(x_i) = X]$, or $f(x_i)$ is the probability that x_i equals a specific value X . Expected values have the following properties:

$$E(A + B) = E(A) + E(B)$$

$$cE(A) = E(cA)$$

$$E(AB) = E(A)E(B) + \text{cov}(A, B)$$

where $\text{cov}(A, B)$ is the covariance between A and B . The first two conditions are those for linear systems, so expected values of a measurement satisfy the minimum criteria for linearity.

Therefore, if your system is nonlinear, then it is only approximated by linear statistical methods such as ANOVA. The approximation grows worse over long enough times or large enough areas such that nonlinear processes strongly determine the experimental outcome. Even though we conceive of biological systems as being nonlinear, we do not really know how to design or analyze a nonlinear experiment. The mathematics of systems of nonlinear equations is currently the only way to rigorously think about biological systems in which we suspect that interactions between species or components are not proportional solely to the sizes of the individual species population densities or the sizes of the individual components.

Exercise 2.2

1 Verify that, in general, 2×2 matrices do not satisfy the commutative property of multiplication. That is, $\mathbf{AB} \neq \mathbf{BA}$ where

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

Under what circumstances would two 2×2 matrices be commutative?

2 Verify that 2×2 matrices satisfy the associative law of multiplication. That is, $\mathbf{C}(\mathbf{AB}) = (\mathbf{CA})\mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are as above and $\mathbf{C}$ is:

$$\mathbf{C} = \begin{bmatrix} i & j \\ k & l \end{bmatrix}$$

3 Verify that 2×2 matrices satisfy the distributive law, or $\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$ where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are as above.

By doing 2 and 3, you have proven that 2×2 matrices are linear systems.

4 Fill in the first two columns of this table (you will fill in the last three columns after the next section).

Matrix	trace	determinant	λ_1	λ_2	Conditions for stability?	Conditions for oscillations?
$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$						
$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$						
$\begin{bmatrix} a & b \\ c & 0 \end{bmatrix}$						
$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$						

Equilibria, eigenvalues, eigenvectors, and stability

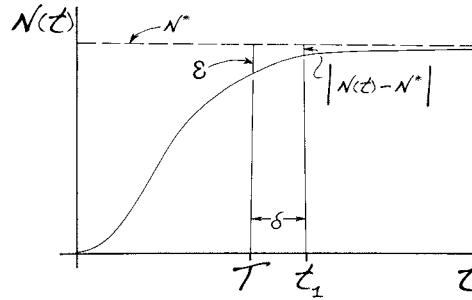
We now have almost all the basic tools needed to construct and analyze a model. The first step in developing a model is to propose an equation or equations for how the population, ecosystem component, or system of interacting components or populations change over time, such as the autonomous differential equation

$$\frac{dN}{dt} = f(N)$$

where N represents the variable of interest (population density, nutrient content, etc.). The time-dependent solution to this equation is $N(t)$ and is found by integrating it with respect to t . Quite often, however, we cannot find the time-dependent solutions by integration, or if such solutions exist they are unwieldy. In such cases, we may wish to focus on their behavior in the neighborhood of key points.

One such set of key points are the *equilibrium* solutions. At equilibrium, the system has a constant value of $N(t)$, for example population density. Call this constant value N^* . Now, Cauchy's definition of a limit can be used to define an equilibrium in the following way: for any arbitrarily small ϵ if we can find a time T such that

Fig. 2-3 The equilibrium of a function is a limit.



$|N(t_1) - N^*| < \varepsilon$ when $t_1 > T$ and the difference between t_1 and T is arbitrarily small (i.e., $< \delta$), then N^* is the limit of $N(t)$ (Fig. 2-3). Put another way, N^* is a constant which is the limit of $N(t)$ as $t \rightarrow \infty$. Since N^* is a constant, then its derivative (rate of change) with respect to t by definition equals 0. So we can find this limit merely by setting dN/dt equal to 0 and solving for any and all N that satisfy this. Such limits are the system's equilibria.

Next, we wish to determine how the system behaves near each equilibrium point. A perturbation of size n near the equilibrium moves the system from N^* to $N(t) = N^* + n$. That is, the perturbation n equals $N(t) - N^*$. The direction and sign of the perturbation depends on whether $N(t)$ is a reduction from N^* ($n < 0$) or an augmentation to N^* ($n > 0$). What we wish to find out is whether the perturbation grows or decays in either direction. If the perturbation grows, then the system moves further away from N^* . N^* is then called an unstable equilibrium. If the perturbation decays, then the system moves back to N^* . N^* is then called a stable equilibrium (Fig. 2-4). How do we determine whether an equilibrium is stable or unstable?

Before developing an algebraic approach to this problem, let's look at a graphical or geometric approach. This will give us an intuitive feel for what is happening which we will then make algebraically more rigorous in a moment. Throughout this book, we will use both geometric and algebraic approaches to problems. You will see the advantages and disadvantages of each approach as we apply them to particular ecological models.

For now, think of $f(N^* + n) = f(N)$ as a value of the differential equation for N within a small neighborhood of some distance n from N^* . We now plot $f(N)$ against N (Fig. 2-5). Note that I have made the function intersect the horizontal axis, or the N -axis, with a negative slope. Since $f(N) = dN/dt$, the point where dN/dt crosses the N -axis and equals 0 is an equilibrium point N^* . Notice that dN/dt is negative within a small neighborhood greater than N^* . Then, over time, the magnitude of

Fig. 2-4 Perturbations decay back to stable equilibria but diverge away from unstable equilibria. (Adapted from figure A.III.6 of Gutierrez 1996.)

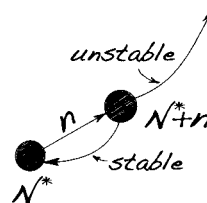
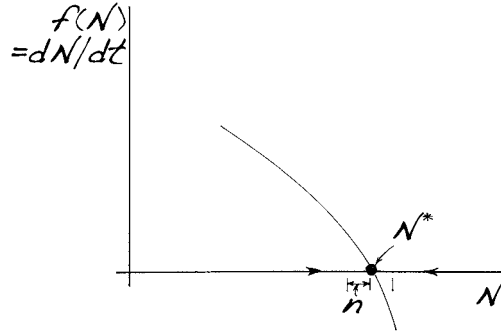


Fig. 2-5 The slope of dn/dt is negative to either side of stable equilibria.



N decreases which means that n also decreases and approaches 0. Therefore, in the limit as $t \rightarrow \infty$, $N \rightarrow N^*$ from the right. On the other hand, dN/dt is positive within a small neighborhood less than N^* . Then, over time, the magnitude of N increases which means that n must once again decrease and approach 0. Again, in the limit as $t \rightarrow \infty$, $N \rightarrow N^*$ from the left.

Let's now put an arrow alongside the N -axis pointing towards N^* from both sides to denote that for any N in this neighborhood, N will be closer to N^* as t increases. Both arrows will point towards N^* because in both cases $N \rightarrow N^*$ as $n \rightarrow 0$. Therefore, since the system returns to N^* from perturbations of size n in either direction, N^* is a stable equilibrium (Fig. 2-5). A stable equilibrium is denoted in figures such as Fig. 2-5 by a solid dot.

We can state this more precisely and compactly in terms of Cauchy's definition of limits. That is, N^* is a limit to the perturbation $(N^* - n)$ applied at t_0 if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that for every $|N^* - n| < \delta$, then $|N^* - n| < \epsilon$. This is called *Lyapunov stability*.

Now let's use these geometric insights to develop a more precise algebraic (analytic) approach to stability. This approach will tell us not only whether the equilibrium point is stable, but also the rate of return to the equilibrium point and how this rate and the stability depends on the parameter values.

Let's first set up the differential equation for the perturbed value of $N(t) = N^* + n$:

$$f(N) = f(N^* + n) = \frac{d(N^* + n)}{dt} = \frac{dN^*}{dt} + \frac{dn}{dt}$$

Since N^* is an equilibrium point, its derivative with respect to time dN^*/dt equals 0, so

$$f(N) = \frac{d(N^* + n)}{dt} = 0 + \frac{dn}{dt}$$

Here, dn/dt is how the perturbation, n , changes with time.

How do we determine the rate of change of the perturbation and how it depends on the model's parameters? To do this, we can make use of the Taylor Series approximation of $f(N)$ which is:

$$f(N) = f(N^* + n) = f(N^*) + n \left. \frac{df(N)}{dN} \right|_{N^*} + \frac{n^2}{2!} \left. \frac{d^2f(N)}{dN^2} \right|_{N^*} + \dots$$

The first two terms are the equation for a straight line which is tangent to $f(N)$ at N^* . This linear function is a good approximation to $f(N)$ if the higher order terms in the Taylor Series approximation are small relative to the first derivative. For most of the equations used in mathematical ecology, this condition is satisfied. We now have

$$f(N) \approx f(N^*) + n \left. \frac{df(N)}{dN} \right|_{N^*}$$

Now $f(N) = dn/dt$ and $f(N^*) = 0$, so

$$\frac{dn}{dt} \cong \left. \frac{df(N)}{dN} \right|_{N^*} n \quad (2.14)$$

This says that the rate of change of a perturbation of size n near N^* is approximately proportional to n with a proportionality constant $df(N)/dN$ evaluated at N^* . The smaller n is, the closer $N(t)$ is to N^* and the slower the change in n . The solution to this equation (which we will derive in more detail later in the chapter on exponential growth) is

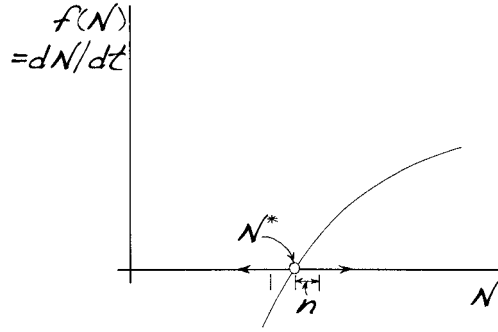
$$n(t) = n(0)e^{\left. \frac{df(n)}{dN} \right|_{N^*} t} = n(0)e^{\lambda t} \quad (2.15)$$

where $n(0)$ is the initial size of the perturbation. It is standard to denote $df(N)/dN|_{N^*}$ more simply by the Greek symbol λ . Since in the above case $df(N)/dN|_{N^*}$ (i.e., the slope of the tangent line at N^*) is negative, the perturbation n will decay exponentially with time. Therefore, for stable equilibria $N^* + n(t)$ will asymptotically approach N^* from either side within a small neighborhood n of N^* where $f(N)$ is approximated by the tangent line at N^* . In this case, N^* is said to be not only Lyapunov stable but *asymptotically stable* with respect to small perturbations.

Now suppose there is another N^* and $f(N)$ has a positive slope near this N^* (Fig. 2-6). By the same reasoning and placing arrows parallel to the N -axis as before we see now that they point away from N^* . Therefore any perturbation of size n near N^* will grow with time because λ is positive. N^* is therefore an unstable equilibrium. An unstable equilibrium is denoted in figures such as Fig. 2-6 by an open dot.

In Eq. 2.15, λ is called the *eigenvalue* evaluated at an equilibrium. Eigenvalues are the instantaneous rates of change of dN/dt with respect to N . When we evaluate the eigenvalues at a particular N^* we denote those with a subscript: λ_{N^*} . If $\lambda_{N^*} < 0$, then N^* is asymptotically stable – that is, the perturbations will decay exponentially with time. If $\lambda_{N^*} > 0$, then N^* is unstable – that is, the perturbations will grow exponentially with time. The arrows on the N -axis are vectors because they point in a particular direction with a magnitude proportional to the rate of change, λ . The collection of arrows showing how N changes for any value of $N(t)$, most especially those values near N^* , is called the vector field of $N(t)$.

Fig. 2-6 The slope of dn/dt is positive to either side of unstable equilibria.



Because the equilibrium is a limit approached as $t \rightarrow \infty$, stable equilibria will never be “reached” in finite time after a perturbation. Nonetheless, using Eq. 2.15, we can estimate the time it would take for the perturbation to decay to some amount less than the magnitude of the initial perturbation, $n(0)$. Conversely, if the equilibrium is unstable, we can also use Eq. 2.15 to estimate the time it would take for the perturbation to grow to some multiple of $n(0)$. Simply define $n(0) = 1$, thus making $n(t)$ some fraction or multiple of $n(0)$. Now, take the natural logarithm of both sides of Eq. 2.15:

$$\ln n = \ln(e^{\lambda t}) = \lambda t$$

and solve for t :

$$t = \frac{\ln n}{\lambda} \quad (2.16)$$

Since we are working to the base e of natural logarithms, it is convenient to define a “natural” time scale of change in the perturbation as some fraction $1/e$ of its original amount (approximately 37%, or a little more than $1/3$). In that case, the characteristic time scale for decay or growth of a perturbation equals $1/\lambda$. We will see later that this defines the turnover rate of a nutrient pool or the mean residence time of a nutrient in a pool.

What if λ_{Ni^*} is a complex number, $a + ib$ (this will actually be fairly common in ecological models)? Recall that:

$$e^{(a+ib)t} = e^{at}e^{ibt} = e^{at} (\cos bt + i \sin bt)$$

The imaginary part of this eigenvalue creates oscillations in the behavior of N because of the correspondence of e^{ibt} with $\cos bt + i \sin bt$. That is, N fluctuates from one side of N^* to the other with a frequency determined by b . The real part of the eigenvalue, a , causes the fluctuations to grow exponentially if $a > 0$ and decay exponentially if $a < 0$. That is, N cycles around N^* either converging to it if N^* is stable or diverging away from it if it is unstable. The sign of the real part of a complex eigenvalue determines the stability whereas the imaginary part produces oscillations.

So in general if the real part of an eigenvalue is negative, the equilibrium is asymptotically stable but if it is positive it is unstable. But the negative sign in the exponent of Eq. 2.15 need not apply to the eigenvalue – it can just as well be associated with t , the time. If time is negative, it is running backward and following the same exponential function only in reverse. Therefore, mathematically we can also distinguish between stable and unstable equilibria with respect to the direction of time in which they are approached. Stable equilibria are approached as $t \rightarrow \infty$ and unstable equilibria are approached as $t \rightarrow -\infty$. We can now define asymptotically stable and unstable equilibria as limits with respect to the direction of time:

$$\begin{aligned} N_S^* &= \{N^* | \lim_{t \rightarrow \infty} (N^* - (N^* + n(0))e^{\lambda t}) = 0\} \\ N_U^* &= \{N^* | \lim_{t \rightarrow -\infty} (N^* - (N^* + n(0))e^{\lambda t}) = 0\} \end{aligned} \quad (2.17)$$

where the subscripts S and U indicate asymptotically stable and unstable equilibria, respectively, and λ is the eigenvalue of dN/dt evaluated at N^* .

What if $\lambda_{N^*} = 0$? To a first approximation, the perturbation will neither grow nor decay with time. Therefore, if we perturb N near N^* by an amount n , it will stay where it is and n will neither grow nor decay. This is called neutral stability. But the key phrase here is “to a first approximation.” Remember that these calculations were done assuming that a tangent line approximates a function near the equilibrium. The higher order terms in the Taylor Series approximation cannot be neglected in such cases. Therefore, the linearization of the behavior of $f(N_i^*)$ is ambiguous when $\lambda_{N^*} = 0$.

For example, a differential equation could be 0 at N^* and be positive to either side of N^* (Fig. 2-7). The arrows show that a perturbation that decreases N to the left of N^* will decay with time and N will approach N^* from the left, but one that increases N to the right of N^* will continue to grow (Fig. 2-7). N^* will then be stable to perturbations that decrease N from it, but not to perturbations that increase N from it. The same thing will happen if the equation is negative to each side of N^* , only this time it will be stable to perturbations to the right but not to the left. N^* in these cases is said to be metastable. Metastable equilibria are denoted by half colored dots, the solid side on the “stable” side and the open side on the “unstable” side.

A differential equation can also be zero at inflexion points (Fig. 2-8). The inflexion point is an equilibrium but whether or not it is stable depends on the sign of the differential equation to either side of the inflexion point.

Fig. 2-7 Perturbations around a metastable equilibrium converge back to it from one side but diverge away from it to the other side.

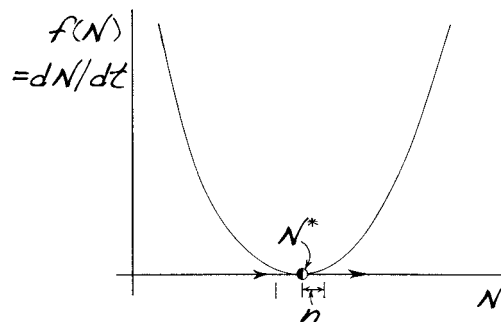
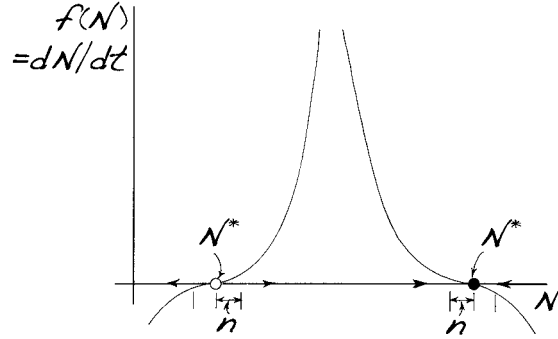


Fig. 2-8 Unstable (left) and stable (right) equilibria at inflexion points of dn/dt .



These methods of stability analysis by calculation of eigenvalues at equilibria and its geometric equivalent of plotting directions of change with perturbations will be the major mathematical tools that we will use to examine the behaviors of ecological models and how they depend on the values of parameters such as birth rate, death and decay rates, predation rates, nutrient inputs and outputs, etc. This extremely valuable tool emerges from thinking about asymptotically stable or unstable equilibria as limits of the solution of a differential equation as $t \rightarrow +/\infty$. Of course, equilibria can and are operationally determined simply by setting the r.h.s. of differential equations equal to 0 and solving for all values of the arguments that satisfy this. But thinking of asymptotically stable or unstable equilibria as limits adds further richness to the concept. First, the system will never get to equilibrium in a finite time interval T either forwards or backwards because the equilibrium is a limit as $t \rightarrow +/\infty$ not as $t \rightarrow T$. Second, as $N \rightarrow N_S^*$, it approaches it at a decreasing rate (or, if you wish, for unstable equilibria in forward time as $N - N_U^*$ increases, it does so at an increasing rate). Therefore, thinking of asymptotically stable or unstable equilibria as limits also tells us something about the behavior of the system away from the equilibria, at least for short distances, whereas thinking of equilibria only as solutions to $dN/dt = 0$ only tells you how to find them, not what the behavior near them is. Finally, the eigenvalues will be functions of the parameters of the differential equation, so we can determine how each parameter (which will correspond to birth, death, decay, uptake, as well as many other rates) affects the stability of the equilibrium. The limit concept therefore adds much more depth and richness to our understanding of equilibria than simply defining them as the 0 solutions of a differential equation.

To examine the stability of a system of coupled differential equations we need to generalize this procedure to a system of k interacting variables, $N_1, N_2, \dots, N_k$, whose dynamics are described by a system of coupled equations:

$$\begin{aligned} dN_1/dt &= f_1(N_1, N_2 \dots N_k; a_1, b_1, \dots) \\ &\vdots \\ dN_k/dt &= f_k(N_k, N_1 \dots N_{k-1}; a_k, b_k, \dots) \end{aligned} \quad (2.18)$$

The trajectory of the system can be described by the flow of a vector whose elements are the N_i . This flow happens in a phase space composed of k orthogonal axes, one axis for each variable. The number of variables and axes of phase space is known as the dimension of the system. The flow of the vector is its evolution over time in this

space as controlled by the coupled differential equations of Eq. 2.18. The equilibrium points are determined by setting the r.h.s. of these equations to 0 and solving for each variable by the usual method of solving simultaneous equations.

To determine the stabilities of the equilibrium points by a Taylor Series expansion for perturbations to each of these differential equations, we must take into account not only the partial derivative of each variable with respect to itself (as in Eq. 2.14 for the single variable case) but also the partial derivatives with respect to all other variables upon which each variable also depends. The linear expansion of a perturbation in k dimensions from equilibrium then becomes:

$$\begin{aligned}\frac{dn_1}{dt} &= \left. \frac{\partial f_1}{\partial N_1} \right|_{N_1^*, \dots, N_k^*} n_1 + \left. \frac{\partial f_1}{\partial N_2} \right|_{N_1^*, \dots, N_k^*} n_2 + \dots + \left. \frac{\partial f_1}{\partial N_k} \right|_{N_1^*, \dots, N_k^*} n_k \\ \frac{dn_2}{dt} &= \left. \frac{\partial f_2}{\partial N_1} \right|_{N_1^*, \dots, N_k^*} n_1 + \left. \frac{\partial f_2}{\partial N_2} \right|_{N_1^*, \dots, N_k^*} n_2 + \dots + \left. \frac{\partial f_2}{\partial N_k} \right|_{N_1^*, \dots, N_k^*} n_k \\ \frac{dn_k}{dt} &= \left. \frac{\partial f_k}{\partial N_1} \right|_{N_1^*, \dots, N_k^*} n_1 + \left. \frac{\partial f_k}{\partial N_2} \right|_{N_1^*, \dots, N_k^*} n_2 + \dots + \left. \frac{\partial f_k}{\partial N_k} \right|_{N_1^*, \dots, N_k^*} n_k\end{aligned}$$

We now have a set of simultaneous linear equations where coefficients of each variable are the partial derivatives. In the previous section on matrices, we saw that a set of simultaneous linear equations can be assembled into a matrix of the coefficients to the variables in each equation times a vector of the variables. Therefore, we can assemble a matrix $\mathbf{J}$ that contains the partial derivatives of each variable with respect to itself in the diagonal elements and the partial derivatives with respect to the other variables in the off-diagonal elements, each evaluated at the equilibrium points of the variables:

$$\mathbf{J} = \begin{bmatrix} \left. \frac{\partial f_1}{\partial N_1} \right|_{N_1^*, \dots, N_k^*} & \left. \frac{\partial f_1}{\partial N_2} \right|_{N_1^*, \dots, N_k^*} & \dots & \left. \frac{\partial f_1}{\partial N_k} \right|_{N_1^*, \dots, N_k^*} \\ \left. \frac{\partial f_2}{\partial N_1} \right|_{N_1^*, \dots, N_k^*} & \left. \frac{\partial f_2}{\partial N_2} \right|_{N_1^*, \dots, N_k^*} & \dots & \left. \frac{\partial f_2}{\partial N_k} \right|_{N_1^*, \dots, N_k^*} \\ \vdots & \vdots & \ddots & \vdots \\ \left. \frac{\partial f_k}{\partial N_1} \right|_{N_1^*, \dots, N_k^*} & \left. \frac{\partial f_k}{\partial N_2} \right|_{N_1^*, \dots, N_k^*} & \dots & \left. \frac{\partial f_k}{\partial N_k} \right|_{N_1^*, \dots, N_k^*} \end{bmatrix} \quad (2.19)$$

This matrix is called a *Jacobian* matrix. It expresses how each component of the system, N_i , changes with respect to itself and also to all other components by means of the partial derivatives of their coupled differential equations (Eq. 2.18). It follows that if these derivatives are evaluated at the equilibrium points $N_1^*, N_2^* \dots N_k^*$ then as with Eq. 2.14 for a single variable,

$$\frac{d\mathbf{n}}{dt} \approx \mathbf{J}|_{N^*} \mathbf{n} \quad (2.20)$$

The formal equivalence to Eq. 2.14 for a single variable linear equation is now clear, except that now the change in each N_i must be evaluated not only against the partial derivative of its differential equation with respect to itself (Eq. 2.18) but also with respect to the other element(s) with which it is coupled. The Jacobian matrix J therefore linearizes the set of coupled differential equations specified by Eq. 2.18 at their equilibrium points. The behavior of the system depends solely on the properties of the Jacobian matrix J , especially on its eigenvalues near the equilibrium points of the system.

The eigenvalues of the Jacobian matrix determine whether the system of differential equations specified by Eq. 2.18 is stable or unstable near an equilibrium. Matrices have eigenvalues which satisfy the following:

$$J\mathbf{w} = \lambda\mathbf{w}$$

where λ is an eigenvalue and $\mathbf{w}$ is an associated (right) eigenvector, which is a nonzero column vector. There are also left eigenvectors which are row vectors that satisfy $\mathbf{w}J = \lambda\mathbf{w}$. Eigenvectors are exactly that, vectors. They point in different directions in the phase space. In almost all ecological models, there usually are as many eigenvectors and associated eigenvalues as there are variables of the system, although each eigenvalue does not usually stand for the rate of change of each variable.

How can we calculate these eigenvalues and associated eigenvectors? The above formula can be rearranged to yield:

$$J\mathbf{w} - \lambda\mathbf{w} = 0$$

$$(J - \lambda I)\mathbf{w} = 0$$

This formula has an obvious solution where $\mathbf{w} = 0$. This is known as the “trivial” solution (although as we shall see it does not have a trivial biological meaning). But we also wish to determine if there are other solutions for $\mathbf{w}$ and associated λ . These other solutions exist if the determinant of the term inside the parentheses equals 0. This rule is our procedure for finding eigenvalues and associated nontrivial eigenvectors of systems of equations:

$$|J - \lambda I| = 0 \tag{2.21}$$

We simply solve for all λ that satisfy this. For example, for a 2×2 matrix:

$$\left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0$$

and using the rule for finding the determinant of a 2×2 matrix,

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{21}a_{12}) = 0$$

This is a quadratic equation of the form

$$a\lambda^2 + b\lambda + c = 0$$

which is a polynomial of degree 2 whose two roots (λ) are guaranteed by the Fundamental Theorem of Algebra to lie in the complex number domain, as discussed above. In this case, we can solve for both λ using the well known quadratic formula

$$\lambda_j = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where $a = 1$, $b = -(a_{11} + a_{22})$, and $c = (a_{11}a_{22} - a_{21}a_{12})$.

The eigenvalues of any matrix are also related to its trace and determinant in several ways:

$$\text{tr} \mathbf{A} = \sum_{j=1}^n \lambda_j \quad (2.22a)$$

$$\det \mathbf{A} = \lambda_1 \lambda_2 \lambda_3 \cdots \lambda_n \quad (2.22b)$$

The eigenvalues of a 2×2 matrix can be found from the trace and determinant by:

$$\lambda_j = \frac{\text{tr} \pm \sqrt{\text{tr}^2 - 4\det}}{2} \quad (2.23)$$

We can now find the eigenvector associated with a particular eigenvalue by using the general definition of an eigenvalue of a matrix given above:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = \lambda \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}$$

and using the elements of the matrix times a vector to determine an equivalent set of simultaneous linear equations:

$$a_{11}N_1 + a_{12}N_2 = \lambda N_1$$

$$a_{21}N_1 + a_{22}N_2 = \lambda N_2$$

which we solve in the usual way for two equations in two unknowns. It turns out that there are an infinite number of pairs x_1, x_2 that satisfy this, each some multiple of a base vector. We find the base vector by dividing each element by the largest, thus expressing the elements of the eigenvector as a fraction of the largest element.

Moreover, if $\det \mathbf{A} \neq 0$ and $\lambda_1 \dots \lambda_k$ are distinct (i.e., no two are equal) and $\mathbf{w}_1 \dots \mathbf{w}_k$ are the eigenvectors associated with each, then $\mathbf{w}_1 \dots \mathbf{w}_k$ are linearly independent (see Hohn 1973, p. 381, Th. 9.3.1 for proof). This implies that we can express any vector $\mathbf{n}(0)$ as:

$$\mathbf{n}(0) = c_1 \mathbf{w}_1 + c_2 \mathbf{w}_2 + \cdots c_k \mathbf{w}_k = \sum_{i=1}^k c_i \mathbf{w}_i \quad (2.24)$$

where the $\mathbf{w}_i$ are the k eigenvectors and c_i is a set of coefficients that rescales the eigenvectors to the desired vector.

Now, just as Eq. 2.15 is an exponential solution for the time trajectory of a perturbation of size $n(0)$ to a single variable (Eq. 2.14), the eigenvalues of the Jacobian matrix express the rate of change of perturbations of size $n_i(0)$ near the equilibrium point in the phase space. These rates of change happen along the eigenvectors which pass through the equilibrium point. This yields the following equation for the change in a perturbation as a function of the eigenvectors and eigenvalues at an equilibrium point:

$$\mathbf{n}(t) = \mathbf{n}(0)e^{\lambda_1 t} = \sum_{i=1}^k c_i \mathbf{w}_i e^{\lambda_i t} = c_1 \begin{bmatrix} w_{11} \\ \vdots \\ w_{k1} \end{bmatrix} e^{\lambda_1 t} + c_2 \begin{bmatrix} w_{12} \\ \vdots \\ w_{k2} \end{bmatrix} e^{\lambda_2 t} + \cdots + c_k \begin{bmatrix} w_{1k} \\ \vdots \\ w_{kk} \end{bmatrix} e^{\lambda_k t} \quad (2.25)$$

Just as the single eigenvalue λ determined the rate of change towards or away from the equilibrium point N^* in the single direction denoted by the N -axis for a single variable, autonomous differential equation, the set of k eigenvalues of the Jacobian matrix determine the rate of change towards or away from the equilibrium N^* in the k directions of the phase space determined by their respective eigenvectors.

For example, let us construct a phase space of two coupled variables N_1 and N_2 with an equilibrium point $N^* = (N_1^*, N_2^*)$. Let the eigenvectors lie orthogonally to each other but obliquely to each axis. Just as above, when we drew arrows indicating the direction of change along the N axis for models of one variable (one dimension) according to whether the eigenvalue was positive or negative, here we draw arrows along the eigenvectors in directions determined by their associated eigenvalues. The “velocity” of change is greatest in the eigendirection associated with the eigenvalue of greatest absolute magnitude ($|\lambda|$). If the arrows along all eigenvectors point towards stable equilibria (all eigenvalues < 0) or away from unstable equilibria (all eigenvalues > 0) in all directions, then such equilibria are called stable or unstable nodes (Figs 2-9, 2-10).

Fig. 2-9 Perturbations around a stable equilibrium converge to it along both eigenvectors.

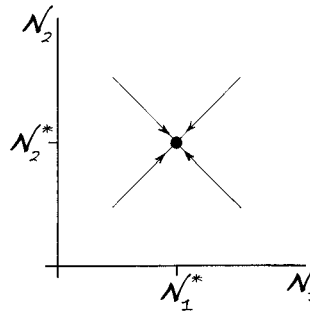


Fig. 2-10 Perturbations around an unstable equilibrium diverge from it along both eigenvectors.

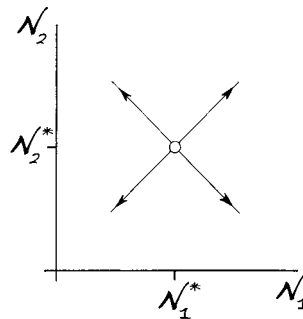
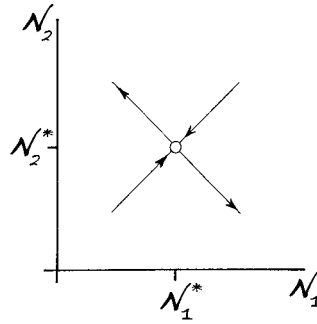


Fig. 2-11 Perturbations around a saddle node diverge from it along one eigenvector but converge to it along the other.



If one eigenvalue is positive and the other negative, then perturbations will be stable only if they are in exactly the direction of the eigenvector associated with the negative eigenvalue. Such an equilibrium point is called a saddle point or saddle node (Fig. 2-11). Imagine you are riding a horse – you are stable to perturbations that rock you in the saddle if the perturbation throws you forward or backward but unstable to perturbations that throw you from side to side. Strictly speaking, a saddle node is an equilibrium which is a limit in forward or backward time only to perturbations exactly along the eigenvectors. Perturbations in all other directions grow as a vector sum of the two eigenvectors: in either forward or backward time, the perturbation will not asymptotically approach the saddle node.

If the eigenvalues are complex conjugate pairs, the imaginary part of e^{a+ib} causes the trajectory to move in a circle but the real part either pushes the trajectory out if $a > 0$ or pulls the trajectory in towards the equilibrium point if $a < 0$. The two motions concurrently cause a spiral trajectory which is either unstable or stable, respectively (Figs 2-12, 2-13). This can be easily seen if we rewrite e^{a+ib} as $e^a e^{ib}$. The term

Fig. 2-12 Complex eigenvalues with positive real parts cause perturbations to spiral away from the equilibrium.

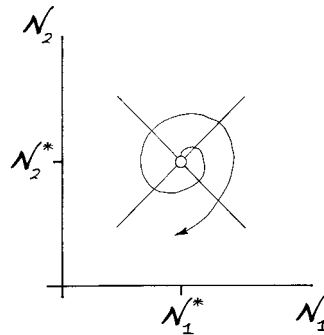
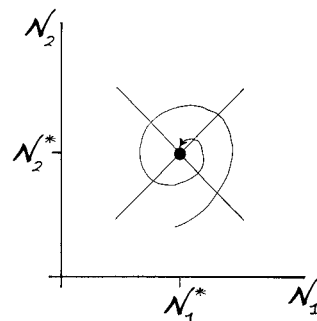


Fig. 2-13 Complex eigenvalues with negative real parts cause perturbations to spiral towards the equilibrium.



e^a causes the trajectory to decay exponentially towards the equilibrium ($a < 0$) or grow exponentially away from the equilibrium ($a > 0$). The term e^{ib} causes the trajectory to spiral around the equilibrium because it equals $\cos b + i \sin b$.

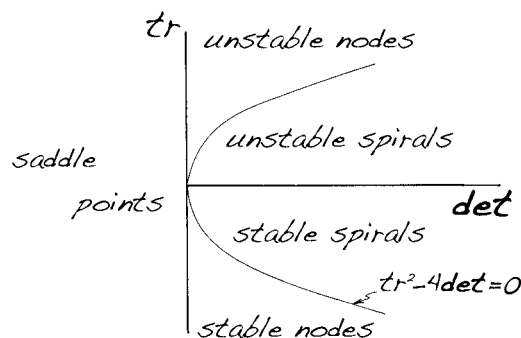
If the real part of complex conjugate eigenvalues equals zero, then trajectories nearby the equilibrium point neither spiral in nor spiral away from it, but just go around in circles concentric to the equilibrium point. This is called neutral stability. Such an equilibrium is called a center and we shall see a classic example in a famous predator–prey equation later in the book. Centers lie on the borderline between stable and unstable spirals. They are very sensitive to the higher order derivatives in the Taylor Series approximation which we have been neglecting in linearizations. Therefore, if the linear approximation of a function has an equilibrium which is a center, it is not necessarily the case that the equilibrium of the nonlinear function is also a center. In that case, you also need to examine the signs of the higher order derivatives as well, not just the sign of the first derivative. On the other hand, if the equilibria are saddles, stable or unstable nodes, or spirals or in the linear approximation, then they really are also saddles, stable or unstable nodes, or spirals near the equilibrium of the full nonlinear system as well.

Directions and magnitudes of change of a perturbation from the equilibrium to any point in the phase space which lies off the eigenvectors are obtained by the vector sum of the directions and rates of change along the eigenvectors in accordance with Eq. 2.25. Eigenvalues and eigenvectors therefore say something about stability in different directions in k -dimensional space.

Finally, we can classify the stabilities of all equilibria of 2×2 matrices (such as those that model two-species interactions) using Eq. 2.21 (Fig. 2-14). To make the real parts of all eigenvalues at an equilibrium < 0 and thus stable, the trace of the matrix must be < 0 . In addition, the determinant must be > 0 . Determinants < 0 yield saddle nodes. The term $tr^2 - 4det$ under the square root sign of Eq. 2.21 defines a parabola. If $tr^2 - 4det < 0$, then the eigenvalue is complex and perturbations will spiral outwards if $tr > 0$ and inwards if $tr < 0$. However, if $tr^2 - 4det > 0$, then perturbations decay exponentially to the stable equilibrium if $tr < 0$ and grow exponentially from the unstable equilibrium if $tr > 0$. Equilibrium points whose eigenvalues have $tr = 0$ are centers. Equilibrium points whose eigenvalues lie on the tr -axis or the parabola are degenerate cases that, while mathematically interesting, will generally not be a property of ecological models.

It is always important to bear in mind that the eigenvectors are straight lines as in the above figures only near the equilibrium points. This straight line character is

Fig. 2-14 Classification of equilibria for two-species models in terms of the trace and determinant of the Jacobian.



a direct consequence of the linearization of the system of differential equations by means of the Jacobian. Far away from the equilibrium point, the nonlinearities of the differential equations may cause the eigenvectors to bend. This bending, to the degree it happens, then complicates the trajectories of the solutions of the equations. Strictly speaking, this technique of stability analysis only applies near the equilibrium points where the linear approximation is a reasonably accurate one. Nonetheless, when such analyses are applied to each of the equilibrium points, we can gain a good qualitative understanding of the flow of the trajectories in the entire phase plane.

Exercise 2.3

Calculate the eigenvalues for the matrices in the table you used for Exercise 2.2, part 4 and determine the conditions required for each eigenvalue to be less than zero and therefore for a stable equilibrium to exist. Then determine the conditions for the eigenvalues to be complex, yielding oscillations.

Summary: what have we learned?

We now have the basic mathematical objects and concepts with which we can explore mathematical models of ecological systems. These include: a set of numbers which are closed under all arithmetic operations that will stand for biological processes; a way to deal rigorously with rates of change (the calculus), including the concept of limits used not only to define instantaneous rates of change but also the existence and nature of equilibria; matrices that allows us to deal with whole properties of sets of coupled equations which will stand for systems of interacting parts; and eigenvalues and eigenvectors, which allow us to assess the stability of an equilibrium point and the behavior of points nearby them.

Mathematics and mathematical notation are extremely abstract – along with perhaps composing music, mathematics is the most abstract form of all thinking. Mathematical ideas and notations are also highly compressed and compact. That is why they are so general. In fact, one of the strategies of doing mathematics is to state an idea or notation as compactly as possible in order to make it more precise and, seemingly paradoxically, more general at the same time. In the process, mathematicians often seem overly cautious, even to other scientists. There is the old joke of a physicist, an ecologist, and a mathematician hiking through Scotland and seeing a black sheep in a field. “Look”, says the physicist (thinking of electrons, which are all alike) “Scottish sheep are black.” The ecologist (who understands the importance of variation of traits) says, “No, some sheep in Scotland are black.” The mathematician (trying to be precise) says “In Scotland, there exists at least one field, containing at least one sheep, which is black on at least one side.”

While precision of thought is the desired end result of mathematics, one of its deep dark secrets is that doing it is not as clean and neat as the finished product seems to be. There is a tension in mathematics between intuitive thinking which constructs new concepts (like Newton did with derivatives) and the hard work required to make them precise (like Cauchy did with limits in response to Berkeley’s criticism). On the one hand, “Constructive thinking, guided by intuition, is the true source of mathematical dynamics . . . The constructive intuition of the mathematician brings to mathematics a nondeductive and irrational element which makes it comparable to music and art” (Courant and Robbins 1961). On the other hand, the hard work of making things more precise is also a necessary part of mathematics

especially because in the process new and very rich sets of properties are often uncovered, such as we have seen with the development of the concept of limits and its application to equilibria and their properties. The payoff is that this rich set of interlocking and precise mathematical concepts helps clarify our thinking about the natural world.

The difficulty of learning or using mathematics comes about when one must “unpack” the compressed, but precisely stated, ideas and notations. This unpacking almost always takes a great deal of thought, time, and hard work. In the remainder of the book we will unpack the ideas of this chapter and develop them further as needed by putting biological “flesh” on them (pun intended!). If these concepts seem abstract to you right now, do what most mathematicians do: don’t worry, just read on in the hopes that things will clear up. Don’t think that you need to understand every point completely before proceeding to the next point. Instead, try to get an overview of the “landscape” first, then go back and fill in the details. What at first seems difficult may become clear in retrospect once you see where the train of thought leads (this usually applies to the author as well!). As the ideas in this chapter become clearer as they gain biological flesh, go back and reread the pertinent parts.

Or read some other books. As you can see, calculus and limits forms a large part of the techniques of mathematical ecology. In recent decades, there has been a push for “Calculus Reform” to find new and (it is hoped) better ways to teach calculus (Gantner 2001), the result of which has been a small cottage industry of producing new calculus texts from both purely mathematical as well as applied perspectives. As a consequence, anyone can find a calculus text to suit them these days. But in my view, you can’t go wrong with the monumental and eminently readable two volume text on calculus and analysis by Courant and John (1965, 1974), which discusses most of the topics in this chapter in greater depth and with admirable clarity without sacrificing rigor. Courant and Robbins (1961) and Klein (1932, reprinted 2004) are also classics that belong on everyone’s reading list. In fact, any book written either by Courant or by Klein is worth reading. Neuhauser (2003) is an excellent recent text on calculus, including eigenvalues, with specific reference to the life sciences by a well-known mathematician and theoretical ecologist. Hight (1977) is an in-depth examination of limits in all their guises. Searle (1966) is the standard reference on applications of matrices to biological problems, but Pettofrezzo (1978) is a succinct and very readable introduction to matrices and their properties. Caswell (2000) is a more advanced treatment of the mathematics of matrix models applied to populations. The MathWorld website (<http://mathworld.wolfram.com/>) gives rigorous thumbnail sketches of concepts in all fields of mathematics and includes some very good examples.

Learning mathematics through its history is an excellent way to come to grips with the subtleties and difficulties of many mathematical objects and operations. The standard scholarly reference on the history of mathematics is the three volume study by Kline (1972), but Kline (1953, 1986) combine history of mathematics with biography and philosophy in a very accessible way. Davis and Hersh (1981) is a set of short but informative, sometimes quirky, but always thought-provoking essays on mathematics as a human activity and the nature of some of its classical problems. Dunham (1990, 2004) presents historical accounts of great theorems

in calculus and other branches of mathematics by showing the way they were originally derived and the difficulties they addressed. The essays in Stewart (1992, 1995) are enjoyable accounts of recent developments in mathematics, including some applications in ecology and the life sciences. The University of St. Andrews website on the history of mathematics is also worth consulting (<http://www-history.mcs.st-andrews.ac.uk/history/>).

Part 2

Populations

3 Homogeneous populations: exponential and geometric growth and decay

Continuous growth and decay (exponential model)

We are now ready to apply the mathematical ideas of Chapter 2 to some ecological problems. Let's start with the simplest case of a population in which birth and death occur continuously and there is a constant per capita birth rate and death rate amongst all individuals of a population. The model we are to develop could be an analog of a population of bacteria in a Petri dish, for example.

The total number of births occurring at any time is simply the per capita birth rate, b , times the population density, N , and the total number of deaths is likewise the per capita death rate, d , times the population density. The change in the population at any one time is the number of births minus the number of deaths, or:

$$\frac{dN}{dt} = bN - dN \quad (3.1)$$

and if $(b - d) = r$, the net per capita change in population density, then

$$\frac{dN}{dt} = rN \quad (3.2)$$

This equation means that something (N) changes at a rate proportional to its size, and the proportionality constant (r) is the instantaneous rate of change. Note that this equation is a linear equation which is in the form $y = mx$, where $y = dN/dt$, $m = r$, and $x = N$. The importance of this will become apparent in a little while.

Solving this simplest of differential equations means finding a function for N at all times such that the slope at all points is rN . To do this, we algebraically separate the variables N and t so that they are on opposite sides of the equation, then integrate both sides:

$$\begin{aligned} \frac{dN}{N} &= rdt \\ \int \frac{dN}{N} &= r \int dt \\ \ln N &= rt + C \end{aligned}$$

where C is the constant of integration. This constant is evaluated by specifying the initial conditions: when $t = 0$, $N = N(0)$, or

$$\ln N(0) = r \times 0 + C$$

$$\ln N(0) = C$$

Substituting $\ln N_0$ back into the above for C we get:

$$\ln N(t) = rt + \ln N(0)$$

$$\ln N(t) - \ln N(0) = rt$$

$$\ln \left(\frac{N(t)}{N(0)} \right) = rt$$

$$e^{\ln \left(\frac{N}{N(0)} \right)} = e^{rt}$$

$$\frac{N(t)}{N(0)} = e^{rt}$$

and solving for $N(t)$ we get

$$N(t) = N(0)e^{rt} = N(0)e^{(b-d)t} \quad (3.3)$$

Now, let's find out whether there are any equilibria of this equation and examine the stability near them by calculating the eigenvalues using the techniques outlined in Chapter 2. To find the equilibria we set the original differential equation equal to 0 and solving for all N that satisfy this. There is only one equilibrium solution to this model:

$$\frac{dN}{dt} = 0 = rN^*$$

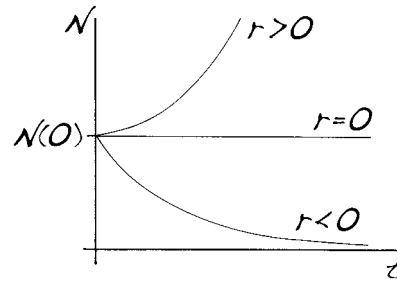
and that is $N^* = 0$. What is the stability of this equilibrium and how does it depend on r ? To determine this, we calculate the expression for the eigenvalue of this model and evaluate it at $N^* = 0$. To find the eigenvalue, we take the derivative of $f(N) = dN/dt = rN$ with respect to N :

$$\frac{d(rN)}{dN} = r = \lambda \quad (3.4)$$

There is only one eigenvalue and it is a constant – the parameter r . Any perturbation of any size n away from $N^* = 0$ will decrease to 0 if $b < d$ ($r < 0$) or increase unbounded and diverge to infinity if $b > d$ ($r > 0$) (Fig. 3-1). Therefore, 0 is an unstable equilibrium point when $r > 0$ and a stable equilibrium when $r < 0$.

It is significant that we don't need to evaluate the eigenvalue at N^* or any other N because it does not depend on N at all. This is because the differential equation is a linear equation and N has dropped out after we took the derivative with respect to it. Linear models are therefore *globally stable* or *globally unstable*. That is, if the model is globally stable, $N(t)$ will converge to the single equilibrium point N^* from

Fig. 3-1 Growth or decay of the exponential model, depending on the sign of r .



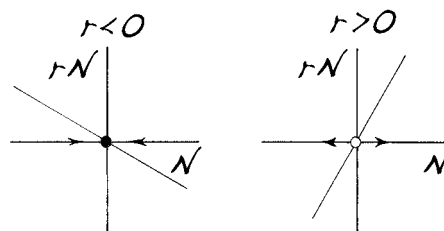
any starting value of N . If the model is globally unstable, $N(t)$ will be repelled away from the single unstable equilibrium point N^* for any value of N .

Later, we will see that nonlinear models, which typically have powers of N , can have several equilibrium solutions. Because these models are in powers of N , taking the derivative with respect to N and evaluating it at an equilibrium point results in an expression for the eigenvalue which is itself a function of N^* . For nonlinear models, the sign of the eigenvalue and the stability of the equilibrium point will depend on which equilibrium point they are evaluated at. We will see that, unlike globally stable or unstable linear models, nonlinear models are typically stable or unstable only within certain local values of N in the neighborhood of a particular N^* .

What happens as r passes through 0 when $b = d$? At $r = 0$ there is a sudden *qualitative* change in the behavior of the model: the equilibrium goes from being unstable to stable because the eigenvalue changes from increasing to decreasing (Fig. 3-2). Such a sudden qualitative change in model behavior as a parameter passes through a critical value is known as a *bifurcation*. Calculating bifurcations, classifying their types, and analyzing their dependence on parameters by means of how eigenvalues near equilibria change sign or change from real to complex as critical values of parameters are passed will be an important technique used throughout this book. This behavior of the exponential model as r passes through a critical value of 0 is the simplest example.

More generally, an equilibrium is said to bifurcate if there are two or more equilibrium solutions that approach it as some parameter a_1 or some combination of parameters $g(a_1, a_2, \dots, a_k)$ approaches a critical value. As the parameter value crosses this critical value, several things may happen: (i) the sign of a real eigenvalue or that of a real part of a complex eigenvalue changes; (ii) the eigenvalue changes from real to complex or vice versa; or else (iii) the eigenvalue changes from complex to purely imaginary ($\text{Re}(\lambda) = 0$). If the sign of the real part of the eigenvalue changes or if $\text{Re}(\lambda) = 0$ at the critical value of the parameter(s), the stability of the equilibrium also changes. A bifurcation analysis has four steps:

Fig. 3-2 Bifurcation in the exponential model as r passes through 0.



- 1 Determine the equilibrium points of the model (or some other point of interest).
- 2 Determine the behavior of the model near the equilibrium point by calculating the eigenvalues at the equilibrium points as functions of the parameters of the model.
- 3 Determine how changing the parameters changes the real parts of the eigenvalues to or from positive to negative, or changes eigenvalues to or from real numbers to complex numbers or to purely imaginary numbers. These are the critical values of the parameters.
- 4 Determine how many solutions bifurcate from the equilibrium point as the parameters pass through their critical points.

In terms of the exponential model, the function $f(N) = dN/dt = rN$ has one equilibrium point, $N^* = 0$ (step 1). The behavior near this equilibrium is given by $df(N)/dN|_{N^*} = r$ (step 2). The value $r = 0$ is a critical value where the behavior changes from increasing away from N^* to decreasing towards N^* (step 3). Therefore, we say there are two solutions bifurcating away from N^* depending on the sign of r (step 4):

Unstable exponential growth: $N(t) = N(0)e^{rt}$, $r > 0$

Stable exponential decay: $N(t) = N(0)e^{rt}$, $r < 0$

Recall that asymptotically stable or unstable equilibrium points are limits. A differential equation has a limit N^* for any t and T if we can find some $\delta > |T - t|$ such that the difference between $N(T)$ and N^* is smaller than any arbitrary value ε . When $r > 0$ and therefore the eigenvalue is positive, $N(T)$ will always be greater than $N(t)$ and therefore farther from $N^* = 0$ when $T > t$. On the other hand, if $T < t$ and we go backward in time $N(T) < N(t)$ and therefore approaches $N^* = 0$. This fits the definition of an asymptotically unstable equilibrium as a limit approached as $t \rightarrow -\infty$ (Eq. 2.16). Conversely, when $r < 0$, $N(T) < N(t)$ for any $T > t$ and $N^* = 0$ is approached in forward time as $t \rightarrow \infty$. Therefore, $N^* = 0$ fits the definition of asymptotically stable equilibrium in Eq. 2.16.

Exercise 3.1

Hall (1964) is an excellent paper that uses exponential decay models to analyze population growth rates of *Daphnia* in Base Line Lake, Michigan in relation to season, temperature, and water chemistry. Hall counted the approximate number of *Daphnia* in 100 liters of water for the following time intervals (data taken from his figure 6):

Time interval	Beginning number of <i>Daphnia</i> $N(0)$	Ending number of <i>Daphnia</i> $N(t)$
January–March	1000	150
April–May	150	3000
June–August	3000	100
September–November	100	2100

Solve the exponential growth equation

$$N(t) = N(0)e^{rt}$$

for r by taking the natural logarithm of both sides. Then use this formula for r to calculate the net rate of population increase or decrease per day during each time interval. What are the relations between birth and death rates for the different time intervals? What hypotheses can you offer to explain the seasonal changes in r ?

Because it is unstable and population size grows unbounded when $r > 0$, the exponential model is not an acceptable model for long-term population growth, although for short terms during which it can reasonably be assumed that r is constant, the exponential model is a reasonable first approximation.

There are, however, several ecological situations for which the exponential model is a very good first approximation for long times when $r < 0$ and it is therefore stable. These situations are when some property decays over time or space. For example, consider a dead leaf which has fallen to the ground. Since carbon is very nearly 50% of the total dry biomass for any living (or recently dead) substance, the dynamics of carbon dominant the decay of organic matter. The leaf has already been “born” at some time in the past, and the carbon atoms (C) in the leaf are not giving birth to any additional carbon atoms. Therefore $b = 0$ at t_0 when the leaf hits the ground. But microbial respiration causes losses or outputs from the leaf’s population of carbon atoms through conversion of organic carbon to CO_2 or for certain anaerobic conditions to CH_4 . This is analogous to the death rate term, d , because both respiration and death cause outputs from the population of interest (carbon atoms in the leaf or number of individuals in the population, respectively). Therefore, the exponential decay model (in which the exponent is usually symbolized by k for “decay rate”) is the starting point for all decomposition models (Olson 1963):

$$C(t) = C(0)e^{-kt} \quad (3.5)$$

This identical model is also used for decay of radioactive substances. The units of k are fraction of original material lost yr^{-1} , or sometimes abbreviated to simply yr^{-1} . Numerous decomposition experiments show that the instantaneous decay rate, k , differs widely for different species’ leaf litters but that there is much less variation in the decay rates of wood and roots (Berg and McClaugherty 2003).

Because $C(t) \rightarrow 0$ only as $t \rightarrow \infty$, $C(t)$ will never “reach” 0 in finite time. But $C(t)$ will reach any other value less than $C(0)$ in finite time. Usually it is interesting to determine the time it takes for $C(t)$ to become some fraction, say half, of $C(0)$. This is easily determined, for example by dividing Eq. 3.5 by $C(0)$ and solving for t :

$$\frac{C(t)}{C(0)} = 0.5 = e^{-kt}$$

or

$$-\frac{\ln(0.5)}{k} = \frac{\ln(2)}{k} = \frac{0.693147}{k} = t \quad (3.6)$$

This is known as the “half life” of decay, and applies equally well to radioactive isotopes as to litter decay.

The fraction $C(t)/C(0) = e^{-kt}$ is the probability that a given carbon atom will reach time t before release by microbial respiration. The mean residence time, R , of a carbon atom in organic form during decomposition is thus the integral over ages 0 to ∞ , or:

$$R = \int_0^{\infty} e^{-kt} dt = \frac{1}{k} \quad (3.7)$$

Therefore while k is the instantaneous loss rate of carbon through respiration, its reciprocal gives the mean residence time of carbon in the sample.

The stability of $C^* = 0$ for exponential decomposition means that all the carbon will eventually be lost from the leaf as $t \rightarrow \infty$. There will be no residual pool of carbon that does not decay. Therefore, this model suggests that there will be no permanent storehouse of carbon in dead material on the earth's surface which will not eventually be returned to the atmosphere as CO_2 or (in anaerobic environments) as CH_4 . Both CO_2 and CH_4 are greenhouse gasses that are responsible for the majority of heat trapped in the atmosphere that makes life possible but that are also increasing (approximately exponentially!) as buried and old organic carbon is mined and burned as fossil fuel. There is much research now on how to remove CO_2 from the atmosphere and store it. A stable equilibrium of $C^* = 0$ implies that to remove CO_2 from the atmosphere, one must sequester carbon from any interaction with microbes who will simply transfer it back to the atmosphere – this is essentially what has happened as dead organic matter during the Carboniferous became buried and turned into coal, oil, or natural gas. Current greenhouse gas remediation policy is to try to sequester carbon either through chemical precipitation of CO_2 from the atmosphere followed by direct burial into ocean sediments (thus mimicking the process of burial that lead to the formation of fossil fuels in the first place) or by managing ecosystems so that carbon taken up in photosynthesis is eventually transferred into substances that decompose slowly, such as long-term storage in wood or in some forms of soil organic matter with slow decay rates.

Exponential decay also happens with respect to increasing distance from an initial point in space rather than increasing from an initial point in time. A good example is the “decay” of light with distance down a water column in a lake or down through a forest canopy. Sunlight is an initial input at the lake surface or the top of the canopy. If light is absorbed at a constant rate with depth, then we have the Lambert–Beers Law:

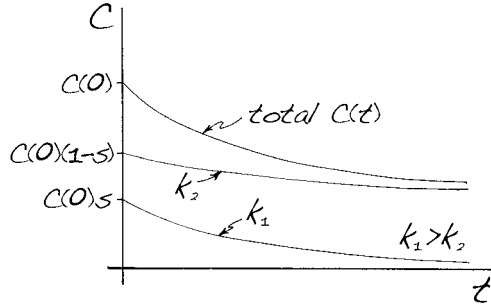
$$I(d) = I(0)e^{-kd} \quad (3.8)$$

where $I(0)$ is incident sunlight, $I(d)$ is amount of sunlight in the same units with depth, d , and k is the light extinction coefficient in units of fraction of incident light absorbed per unit depth. This is the starting point for all models of plant canopy or aquatic ecosystem photosynthesis (Monsi and Saeki 1953 and translated 2005, Nobel et al. 1993).

There are several hidden assumptions to the exponential model that need to be discussed. The first is the assumption of homogeneity. That is, the parameter values for b and d (or for k in exponential decay models) apply to any subdivision of the population in question. For populations, the exponential model assumes that all individuals have the same birth and death rates. We will discuss ways to relax this assumption in population growth models in subsequent chapters.

For decomposition models, it is assumed that all the carbon fractions decay at the same rate. This assumption is easily relaxed by assuming that each carbon compound

Fig. 3-3 Decay as a sum of two independent pools of carbon.



in a leaf has its own intrinsic rate of decay and that the decay of each compound does not affect the decay of another. This leads to exponential decay models that are the sum of decays of at least two fractions (Minderman 1968):

$$C(t) = sC(0)e^{-k_1 t} + (1 - s)C(0)e^{-k_2 t} \quad (3.9)$$

where s is the fraction of $C(0)$ that decays at rate k_1 and $(1 - s)$ is the remaining fraction that decays at rate k_2 , $k_1 \neq k_2$. As decomposition proceeds, the decay rate of the entire substance passes smoothly from being dominated by that of the fraction with larger k to being dominated by that of the fraction with smaller k (Fig. 3-3).

A double exponential decay model appears to fit data from very long term decomposition experiments that run for a decade or more (Berg and McClaugherty 2003). In theory, one could propose as many fractions to this model as there are carbon compounds in a leaf or any other decaying substance, but in practice usually two fractions statistically account for >95% of mass or carbon loss (e.g., Bridgman et al. 1998, Berg and McClaugherty 2003). This has led ecosystem ecologists who work on problems of decomposition to refer to labile and recalcitrant fractions of carbon compounds, the labile fraction having the greater k_i .

Exogenous environmental conditions, such as climate or presence or absence of oxygen, can also affect k . The decay rate k can then be expressed as a function of some hypothesized controlling factor, such as temperature, precipitation, or redox potential (Eh) which may change in time ($k(t) = f(\text{controlling factors}, t)$). That function is then inserted in place of k in Eq. 3.2. Then,

$$\frac{dC}{dt} = k(t)C$$

which has the solution

$$C(t) = C(0)e^{\int k(t) dt}$$

Making parameters in a model (which are constants) functions of some other controlling variables is one way to generalize models so that they describe a wider set of circumstances.

Exercise 3.2

Leaf litter of species X is composed of 25% of compound A which decays at rate $k_1 = 0.5 \text{ yr}^{-1}$ and 75% of compound B which decays at rate $k_2 = 0.1 \text{ yr}^{-1}$. What is the half-life of each compound? How much of the entire compound will remain when half of compound A is left, and how much will remain when half of compound B is left?

The Lambert–Beers Law also assumes homogeneous scattering or absorption of a photon by a leaf with depth in the canopy or by a molecule with depth in the water column. But local increases or decreases in the leaf area in a canopy or the transparency of the water column will alter the absorption or scattering coefficient with depth. Then, one must replace d in Eq. 3.8 with a function describing how leaf area index ($LAI(d)$), leaf area per unit ground area within distance d from the top of the canopy) or water transparency changes with depth and integrate the function down to depth δ :

$$I_\delta = I_0 e^{-k \int_0^\delta LAI(d) dd} \quad (3.10)$$

The units of the extinction coefficient, k , must now be in fraction of incident light absorbed per unit leaf area index or equivalent units of water transparency rather than fraction of light absorbed per unit depth. Monsi and Saeki (1953 and translated 2005) analyze this in more detail and show how to obtain the integral of leaf area index with canopy depth.

Birth and death as discrete events (geometric model)

Another hidden assumption of the exponential model is more subtle. The differential equation which we proposed as Eq. 3.1 is continuous – that is it assumes that the argument (N or C) is infinitely divisible. But the biological world is discrete, whether we are talking about individuals in a live population or individual carbon atoms (Durrett and Levin 1994). No single individual gives birth continuously or has some fraction of itself die continuously – birth and death are discrete events. Therefore, continuity is approximated only when there are a large number of individuals distributed amongst overlapping generations, each with an equal probability of giving birth or dying at any instant. It is the probability of giving birth or dying (or being converted from organic carbon to CO_2) that is homogenous amongst all individuals of a population. When there are discrete, nonoverlapping generations (e.g., in univoltine insects such as mayflies whose males die after mating once and whose females lay one set of eggs per year then die), or where there are small numbers of individuals in the population, the dynamics are better described by discrete or difference equation models. In order to examine more closely growth or decay as discrete processes, let us now derive a model for growth of a population of discrete individuals who reproduce and die in discrete events.

To do this, we must set up the mass balance equation of birth and deaths of discrete individuals as a difference equation:

$$N_{t'+1} = N_{t'} + BN_{t'} - DN_{t'} \quad (3.11)$$

where B is number of births per individual over a single time step, D is probability of death of an individual over a single time step, and $N_{t'}$ is the total number of individuals in the population at time t' (the reason why we are using t' instead of t will become apparent in a moment). Now subtract $N_{t'}$ from both sides then factor out $N_{t'}$ from the r.h.s and we get the discrete analog to Eq. 3.1:

$$N_{t'+1} - N_{t'} = \frac{\Delta N}{\Delta t'} = (B - D)N_{t'} \quad (3.12)$$

(Note: for discrete equations it is customary to use subscripted indices, e.g., $N_{t'}$, rather than parenthetical indices, e.g., $N(t)$, as in differential equations.)

Just as dN/dt in Eq. 3.1 is the limit of $\Delta N/\Delta t'$ as $\Delta t' \rightarrow 0$, so too is $(b - d)N$ the limit of $(B - D)N_{t'}$. And $N^* = 0$ is the equilibrium for both models when $b = d$ and when $B = D$. As was the case with the exponential model, the discrete model is also unstable when $B > D$ and stable when $B < D$. The two linear equations seem nearly identical in several respects, but the behaviors of the two equations are distinctly different in other respects. To see the difference between the two models, start again with Eq. 3.11 only begin by factoring out $N_{t'}$ on the r.h.s.:

$$N_{t'+1} = (1 + B - D)N_{t'}$$

Incidentally, for this equation to have biological meaning, D must be less than $1 + B$, otherwise $N_{t'}$ changes sign at each time step.

Call $1 + B - D$ the net rate of increase, or R , yielding

$$N_{t'+1} = RN_{t'} \quad (3.13)$$

Now start with an initial population size N_0 . The number of individuals after the first time step is:

$$N_1 = RN_0$$

and after the second time step:

$$N_2 = RN_1 = R(RN_0) = R^2N_0$$

and so the general model is:

$$N_{t'} = R^{t'}N_0 \quad (3.14)$$

This discrete equation or map is called the geometric growth model. Now, this doesn't exactly look like Eq. 3.3, the solution of the continuous differential equation model for $N(t)$, and it isn't (Fig. 3-4). The reason is that t' in Eq. 3.14 is measured in discrete (integer-valued) steps whereas t in Eq. 3.3 is continuous and measured in real numbers. So, just as dN/dt in Eq. 3.1 is the limit of $\Delta N/\Delta t'$ as $\Delta t' \rightarrow 0$, so too is $N(0)e^{rt}$ the limit of $R^{t'}N_0$ as $\Delta t' \rightarrow 0$. Having made the point that time in discrete models is not precisely the same as time in continuous models, we will use t in the rest of the book to stand for both, but you should be aware of and sensitive to the difference between the two when switching between continuous and discrete forms of models.

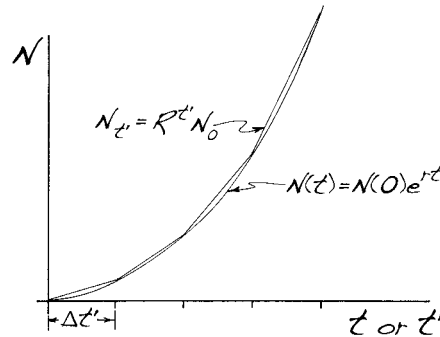
R in the discrete, geometric model performs the same operation as e^r in the continuous, exponential model. The two are related as:

$$R = e^r \quad (3.15a)$$

or

$$r = \ln R \quad (3.15b)$$

Fig. 3-4 The continuous exponential model is the limit of the discrete geometric model.



Even though they both depict the net balance between births and deaths, R and r are not the same thing mathematically. From Eq. 3.14, you can see that there is also a critical point for R that separates unbounded growth way from an unstable equilibrium to decay to asymptotic decay to a stable equilibrium. But, whereas for the exponential model the analogous critical value of r was 0 when $b = d$, here the critical value of R (R_{crit}) equals 1 when $B = D$. When $R > 1$, N_t increases after each time step but when $R < 1$, N_t decreases by the some fraction of the previous time step. And so, we have a similar sort of bifurcation in the geometric model as in the exponential model occurring at a critical value of an analogous parameter, but the critical values for the two models differ. Even in this simplest case, we need to be careful when taking over concepts from continuous differential equation models to discrete difference equation models. When we discuss nonlinear models, differences in the behaviors of continuous and discrete form will become even larger.

The geometric model has a time-honored place in the history of ecology and evolutionary biology. In 1826, Thomas Malthus (1766–1834) pointed out in his famous book that a population grows geometrically whereas the resources to support that population, X , grow arithmetically at best (that is by addition of new resources as they are discovered: $X_{t+1} = X_t + x$, where x is the net amount of new resources discovered during a time step). The importance of Malthus's observation is that it inspired Charles Darwin to begin thinking about how unchecked population growth outstrips the supply of its limiting resources, thereby decreasing reproductive output of those individuals less able to obtain limiting resources. This disparity between geometric and arithmetic growth is the underlying ecological basis of natural selection. This thought lead Darwin on a decades long intellectual journey culminating in the publication of *The Origin of Species* in 1859, which marked the beginnings of modern ecological science.*

*It is of some historical interest to note that Malthus's publisher, John Murray, originally of Edinburgh but in business in London, was also Darwin's publisher as well as that of Michael Faraday, David Livingston, Robert Falcon Scott, and many other British and American scientists, naturalists, and explorers of the nineteenth century, to say nothing of the novels of Jane Austen, Herman Melville, Sir Walter Scott, and Charles Dickens, the poems of Lord Byron, and books by scores of other notable English speaking writers. Many of these books were illustrated by the British artist J. M. W. Turner. The English speaking world owes a great debt to this publishing firm, which unfortunately is no longer in independent business. The National Library of Scotland maintains original manuscripts and letters from many of John Murray's authors in its John Murray Archives (<http://www.nls.uk/jma/>).

Summary: what have we learned?

Even the simplest ecological model, where there is a constant net per capita rate of change of a population (r or R), has a more complicated behavior than at first sight. For these models, the only equilibrium is 0, but whether 0 is unstable or stable depends on whether the per capita rate of change is positive or negative, respectively. As this rate of change passes through a critical value, the equilibrium changes from being unstable to being stable, or vice-versa depending on the direction of change. The qualitative behaviors of the models are said to bifurcate at a critical value of the parameter r or R . But this critical value depends on whether the model is a continuous differential equation ($r_{crit} = 0$) or a discrete difference equation ($R_{crit} = 1$). Even simple linear models of change have underlying complications that urge mathematical caution!

The divergence of the exponential or geometric models to infinity when net growth rate is positive makes them unsuitable for modeling population growth, especially over long time periods. Population ecologists have developed many ways to relax the assumptions of constant per capita birth and death rates in order to overcome this unrealistic divergence. We shall be exploring many of these modifications in the next few chapters.

By contrast, ecosystem ecologists have found the stable form of the exponential or geometric models when net per capita change is negative to be very useful first approximations of decomposition or of capture of light down through a canopy or water column. The exponential or geometric decay model is the basis of all decomposition and canopy photosynthesis models. Further relaxations of the assumption of homogeneity of decay constants among materials or parts of a canopy or between different environments, canopies, or species has led to a rich literature in ecosystem ecology based on elaborations of the exponential decay model.

Open questions and loose ends

Research continues on further mathematical elaborations of exponential decay equations as models of the decomposition process. In particular, while the double exponential decay model for the sum of labile and recalcitrant materials (Eq. 3.9) seems to statistically account for most of the variation of decay rates over time of many types of litter and soil organic matter, there is a conceptual flaw hidden in it. Litter and soil organic matter do not decay simply by having different substances in them being transformed independently to CO_2 or CH_4 during microbial respiration. This very passive view of microbes sees them as simply pass-through agents for carbon from organic matter to the atmosphere. But microbes do more than respire organic carbon, they also convert different forms of carbon into other forms. Almost always, these transformations are accompanied by a degradation of what has become known as the quality of the organic matter: that is, the ability of subsequent populations of microbes to transform or respire the organic matter further decreases (Berg and McClaugherty 2003). Not only are new but more recalcitrant forms of organic matter formed during microbial decomposition, these new forms often physically mask other labile forms, thus making it harder for microbes to subsequently access the labile forms. Even in undecomposed litter in its initial pristine state, recalcitrant compounds such as lignin or cellulose physically mask some proportion of labile forms as sugars, starches, and proteins.

Therefore, despite its empirical success at accounting for the majority of variation in data, Eq. 3.9 is theoretically inadequate because the decay of each type of organic matter is independent of the other and because the creation of new substances

during decomposition is not accounted for. The assumption of the independence of the decay of different substances implicit in Eq. 3.9 must be relaxed. New advances in mathematical modeling of decomposition are likely to come by deriving and analyzing master equations of the decomposition process that are flexible enough to account for transformations of organic matter, especially the progressive loss of quality during decomposition. A very good start in this direction has been made by Ågren and Bosatta (1998), who propose a master equation which models a continuous decline in organic matter quality (a decline in the decay rate k) as decomposition proceeds. The carbon output from the decay of all forms of organic matter is then given by an equation which integrates a distribution function along this quality gradient. An advantage of this approach is that it readily collapses to the simple exponential decay model when conditions of homogeneity of materials and environment are re-imposed. The dynamics of this distribution function and the coupling of this model with a model of carbon and nutrient uptake by the live plant community is a rich area for further research.

4 Age- and stage-structured linear models: relaxing the assumption of population homogeneity

In the previous chapter we saw that exponential and geometric decay models of populations which grow at a constant rate are unstable because the eigenvalue is positive for all $N > 0$, causing the population to grow unbounded. How can this problem be corrected? Perhaps one or more of the assumptions is wrong. One of the assumptions of the exponential and geometric models is that per capita birth and death rates are constant and identical for all members of a population. But this is obviously not the case in nature, where both birth and death rates differ for young, middle-aged, and elderly members of a population or for different life stages of a population such as organisms that undergo metamorphosis or have complicated life cycles. What if we relax the assumption of constant per capita birth and death rates and made them differ amongst age classes or life stages? Would this stabilize the model and prevent the population from growing unbounded?

Age-structured models

Let's begin by developing a model in which per capita birth and death rates depend on integer-valued (discrete) age classes, corresponding usually to years but applicable to any reasonable time step. First, we need a fecundity equation that sums the birth of all individuals of older age classes to give the population density of the youngest age class, which is often known as the age-0 or young-of-the-year age class. The equation for the inputs to the youngest age class is:

$$n_{1,t+1} = b_2 n_{2,t} + b_3 n_{3,t} + \dots + b_k n_{k,t}$$

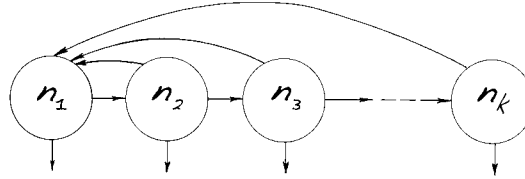
where b_i is the per capita birth rate of the i th age class and there are k age classes. We assume that young-of-the-year do not give birth.

Now we need an equation that gives the survival of individuals of all age classes. If d_i is the probability of an individual of age class i dying, then $s_i = 1 - d_i$ is the probability of an individual of age class i surviving to the next age class. Therefore, the number of individuals of all age classes that survive to the next age class is:

$$n_{i+1} = s_i n_i$$

Note the similarity between the forms of these equations and the geometric growth function. The basic unit of each equation is a per capita rate parameter (birth, death, survival, or net growth) times the density of an age class or, in the case of the geometric model, the entire population. This will be the key to the behavior of the completed model.

Fig. 4-1 Feedbacks between age classes in age-structured models.



This age-structured model has a number of positive feedback loops consisting of younger age classes feeding into older age classes, which in turn supply inputs to the youngest age class (Fig. 4-1). This positive feedback is constrained or “balanced” by the death rates of each age class, which form a sequence of negative feedbacks. The identification of positive and negative feedbacks is an important part of the analysis of models and will be used throughout this book. DeAngelis et al. (1980) is an excellent introduction to positive feedbacks in natural systems.

These two equations can be combined into one matrix:

$$\begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{k-1} & b_k \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & s_{k-1} & 0 \end{bmatrix} \quad (4.1)$$

Now if we multiply this by a vector of the number of individuals in each age class, we get:

$$\begin{bmatrix} 0 & b_2 & b_3 & \cdots & b_{k-1} & b_k \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & s_{k-1} & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ \vdots \\ n_{k-1} \\ n_k \end{bmatrix}_t = \begin{bmatrix} 0 + b_2 n_2 + b_3 n_3 + \cdots + b_{k-1} n_{k-1} + b_k n_k \\ s_1 n_1 + 0 + 0 + \cdots + 0 \\ 0 + s_2 n_2 + 0 + \cdots + 0 \\ 0 + 0 + s_3 n_3 + \cdots + 0 \\ \vdots \\ 0 + 0 + \cdots + s_{k-1} n_{k-1} + 0 \end{bmatrix}_{t+1} \quad (4.2)$$

Note that the first row in the new vector is the fecundity equation for the young-of-the-year age class and the remaining rows are the survival equations to the next age class for all older age classes. Because $b_i > 0$ for at least one age class, the first row represents the positive feedback due to age-dependent birth. The survival probabilities form the subdiagonal of the matrix. Because $s_i < 1$ for all age classes, the subdiagonal represents the negative feedback due to age-dependent mortality (not all members of an age class survive to the next older class). All other elements of the matrix equal 0. We also assume that, by definition, all individuals of the oldest age class die ($s_k = 0$). The proportional age class distribution is the fraction of the population in each age class, and is obtained by dividing each entry in the population age vector by the sum of the entries in the vector. Eq. 4.2 can be written in more compact form:

$$\mathbf{Ln}_t = \mathbf{n}_{t+1} \quad (4.3)$$

where L is the matrix in Eq. 4.1 and $\mathbf{n}$ is the vector of population density amongst age classes.

The matrix in Eq. 4.2 is known as a Leslie matrix and the model in Eq. 4.2 and 4.3 is known as a Leslie model, in honor of Patrick H. Leslie who invented it. Leslie developed matrix population models on the suggestion of Charles Elton, under whom he worked in the Bureau of Animal Population at Oxford. This Bureau was established to do research on population dynamics of wild game and animal pests. During World War II, its primary target of research was the population of rats in the London sewers: the government was worried that the rat population would erupt and spread plague as public health, water supply, and sewage services were disrupted by the bombings, perhaps the first recognized problem in urban ecology. Elton wondered if the fertility and mortality data about an organism, which the Bureau was collecting using the newly developed mark-and-recapture techniques, could be combined into one mathematical expression. Leslie, who had taught himself matrix algebra while previously confined in a hospital with tuberculosis, adapted these techniques for describing age structure of populations and later published them in a paper which is remarkable for its clarity (Leslie 1945). Nonetheless, ecologists ignored his paper for many years because they felt that matrix algebra was too arcane. It was only in the 1970s that his techniques began to be used extensively.

Note the similarity of Eq. 4.3 to the geometric growth model, Eq. 3.13 – in fact, they are in a sense identical. The matrix L in Eq. 4.3 replaces the scalar R in Eq. 3.13 and the vector $\mathbf{n}_t$ in Eq. 4.3 replaces the scalar n_t in Eq. 3.13. Otherwise we are multiplying an object that stands for population density (n_t or $\mathbf{n}_t$) by another object that stands for the net balance between birth and deaths of individuals (R or L).

Because the Leslie model of an age-structured population is equivalent to the geometric model of growth of a homogeneous population, the equilibria of the two models are also equivalent. For the geometric model, $N^* = 0$ and for the Leslie model $\mathbf{n}^* = \mathbf{0}$, where $\mathbf{0}$ is a vector in which all entries = 0. However, the equilibrium condition $\mathbf{n}_{t+1} = \mathbf{n}_t$ can be satisfied by a matrix L in which the number of births over all age classes (given by the first row of L) is exactly matched by the sum of deaths over all age classes (given by the survival probabilities given in the subdiagonal of L). But such a fine balancing of births and deaths is extremely unlikely in nature and constitutes a special case.

What about the stability of the equilibrium of the Leslie model? Recall that the eigenvalues and eigenvectors of a matrix satisfy the condition:

$$L\mathbf{w} = \lambda\mathbf{w}$$

where $\mathbf{w}$ is an eigenvector and λ is an associated eigenvalue of L which satisfies the equation:

$$|L - \lambda I| = 0$$

Because of the formal equivalence of the Leslie and geometric models, the eigenvalue is performing an identical mathematical operation to R – the eigenvalue is the rate at which the Leslie matrix L causes the elements of the associated eigenvector $\mathbf{w}$ to change.

Now, recall the discussion of eigenvalues and eigenvectors for matrices in Chapter 2. Suppose we want to project an initial population vector $\mathbf{n}_0$ forward in discrete steps

by multiplying it by a matrix A . A Leslie matrix has k distinct, real eigenvalues (Caswell 2000). Therefore, we can project the population vector forward using the equivalence of $Lw = \lambda w$:

$$\mathbf{n}_1 = A\mathbf{n}_0 = \sum_{i=1}^k \lambda_i c_i \mathbf{w}_i$$

Where λ_i and $\mathbf{w}_i$ are the distinct eigenvalues and associated eigenvectors and c_i is a rescaling coefficient. For $\mathbf{n}_2$, we obtain:

$$\mathbf{n}_2 = A\mathbf{n}_1 = A(A\mathbf{n}_0) = \sum_{i=1}^k \lambda_i \lambda_i c_i \mathbf{w}_i = \sum_{i=1}^k \lambda_i^2 c_i \mathbf{w}_i$$

and so in general for any iteration t :

$$\mathbf{n}_t = A^t \mathbf{n}_0 = \sum_{i=1}^k \lambda_i^t c_i \mathbf{w}_i \quad (4.4)$$

Note that this is formally equivalent to the equation for geometric growth of a single variable (Eq. 3.14) but now in terms of all the vector elements in $\mathbf{n}$ rescaled by c_i and with R now replaced by all of the λ_i . Thus, the set of eigenvalues of a Leslie matrix is formally equivalent to the intrinsic rate of population change in the geometric model of the growth of a homogenous population. The solution of age-dependent population growth is the weighted sum of geometric growth for k ages, the weighting factors, c_i , determined by the initial conditions. An inspection of this equation yields the following rules of population behavior (Caswell 2000):

- if $\lambda_i > 1$, it causes population density to increase geometrically;
- if $\lambda_i = 1$, then $\lambda_i^t = 1$ (no effect on population density);
- if $0 < \lambda_i < 1$, then it causes population density to decrease geometrically;
- if $-1 < \lambda_i < 0$, then it causes population density to decline damped oscillations;
- if $\lambda_i = -1$, then it causes population density to oscillate undamped;
- if $\lambda_i < -1$, then it causes population density to increase with oscillations whose amplitude increases geometrically. This will inevitably cause some age classes to become negative.

What if λ_i is complex? Complex eigenvalues of a matrix come in complex conjugate pairs, $x + iy$ and $x - iy$. Recall from Chapter 2 that the distance from the origin to any point (x, iy) in the complex plane is (from the Pythagorean Theorem) the square root of $x^2 + y^2$, which is also the square root of the modulus of the complex number:

$$\sqrt{|\lambda|} = \sqrt{(x + iy)(x - iy)} = \sqrt{x^2 + y^2}$$

Raising a complex eigenvalue to a power does two things: it rotates the position of the complex number in the complex plane and, depending on the modulus, it stretches the distance to the origin, shrinks the distance to the origin, or keeps the distance constant. The distance stretches while the position rotates if $|\lambda| > 1$, it shrinks as it rotates if $|\lambda| < 1$, and it simply rotates without stretching or shrinking if $|\lambda| = 1$.

A circle inscribed by rotating a complex number whose modulus equals 1 is called the unit circle. Therefore, complex eigenvalues whose modulus falls outside the unit circle will cause the associated eigenvector to oscillate and grow geometrically. Complex eigenvalues whose modulus falls inside the unit circle will cause the associated eigenvector to oscillate and decay geometrically and complex eigenvalues whose modulus is on the unit circle will always remain on the unit circle and cause the associated eigenvector to oscillate without growing or decaying. Thus, the position of the complex eigenvalue with respect to the unit circle is analogous to the above three rules for real numbers with respect to 1 or -1 .

Exercise 4.1

To illustrate how raising a complex number to a power rotates and either shrinks, does nothing, or stretches the distance from the origin depending on whether $|\lambda|$ is less than 1, equal to 1, or greater than 1, consider these three eigenvalues:

$$\lambda_1 = 0.5 + 0.3i \quad \lambda_2 = 1/\sqrt{2} + 1/\sqrt{2}i \quad \lambda_3 = 1.5 + 1.0i$$

then do the following:

4.1a calculate the modules of each eigenvalue;

4.1b plot the position of each eigenvalue in the complex plane along with the unit circle;

4.1c square each eigenvalue, calculate the modulus of the squared eigenvalue, and repeat step 4.1b. Discuss what happens to the eigenvalues with successive multiplications.

It is sometimes useful to know the period of these oscillations, which is determined by the ratio of the real and imaginary parts of λ_i (Caswell 2000):

$$P_i = \frac{2\pi}{\tan^{-1}\left(\frac{\text{Re}(\lambda_i)}{\text{Im}(\lambda_i)}\right)} \quad (4.5)$$

The above rules can therefore be simplified to:

- if $|\lambda_i| < 1$, it causes population density to decrease geometrically, smoothly if $\lambda_i > 0$, and with damped oscillations if $\lambda_i < 0$ or complex;
- if $|\lambda_i| > 1$, it causes population density to increase geometrically, smoothly if $\lambda_i > 0$, and with oscillations of increasing amplitude if $\lambda_i < 0$ or complex.

What determines the long-term behavior of a population age-structure vector multiplied by a Leslie matrix as $t \rightarrow \infty$? Leslie matrices have a unique, positive, real eigenvalue λ_1 which is greater than $|\lambda_i|$ for all other eigenvalues. Now factor out λ_1^t on the r.h.s of Eq. 4.4:

$$\mathbf{n}_t = \mathbf{A}^t \mathbf{n}_0 = \lambda_1^t \left(c_1 \mathbf{w}_1 + \left(\frac{\lambda_2}{\lambda_1} \right)^t c_2 \mathbf{w}_2 + \left(\frac{\lambda_3}{\lambda_1} \right)^t c_3 \mathbf{w}_3 + \cdots + \left(\frac{\lambda_k}{\lambda_1} \right)^t c_k \mathbf{w}_k \right) \quad (4.6)$$

Now since $|\lambda_i|/\lambda_1 < 1$,

$$\lim_{t \rightarrow \infty} \mathbf{n}_t = \lambda_1^t c_1 \mathbf{w}_1 \quad (4.7)$$

That is, for any t_1 , we can always find a larger t_2 to make the terms in $|\lambda_i||\lambda_1|$ even smaller and thus make $\mathbf{n}_t$ within any specified ε of $\lambda_1^t \mathbf{w}_1$.

The long-term behavior of the population is therefore determined by the largest real eigenvalue, λ_1 , known as the dominant eigenvalue (unless otherwise specified I will follow the convention in mathematical ecology and use λ to refer to the dominant eigenvalue for the remainder of this chapter). The eigenvector associated with the dominant eigenvalue is called the dominant eigenvector. Although the magnitude of each element of the dominant eigenvector is growing geometrically, the proportions or ratios amongst all the elements converge to constant values. These proportions can be determined by dividing each element of the dominant eigenvector by the total population density given by the sum of all the elements. The dominant eigenvector of $\mathbf{L}$ is then the stable proportional age class distribution for any initial population vector. Therefore, although total population density increases without bounds in a Leslie model, the proportional distributions of densities converge on a stable age distribution given by the dominant eigenvector.

The ratios of the other eigenvalues to the dominant eigenvalue in Eq. 4.6 determine the transient behaviors and rates of approach of the population vector to the stable age distribution. The largest of these ratios is, of course, λ_2/λ_1 and this has the greatest effect on the transient behaviors, the effect of the other ratios decreasing proportionally to their difference with this ratio. The magnitude of this ratio, $|\lambda_2/\lambda_1|$, is called the damping ratio and is usually symbolized by ρ (Caswell 2000). The time required for the contribution of λ_1 to become x times as great as that of $|\lambda_2|$ is:

$$\rho^t = \left(\frac{\lambda_2}{\lambda_1} \right)^t = \frac{1}{x}$$

And by taking the logarithm of both sides and solving for t_x we get:

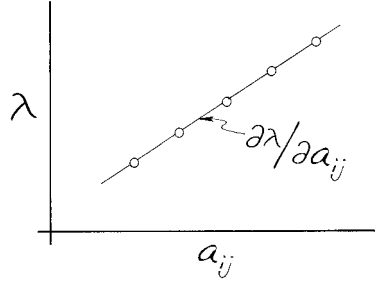
$$t_x = -\frac{\ln x}{\ln \rho} \quad (4.8)$$

(the negative sign does not mean that time is negative but is needed because $\ln \rho$ is negative since $\rho < 1$). The percent of convergence to the stable age distribution is $100 - (100/x)$. For 95% convergence to the stable age distribution, t is solved for $x = 20$ (i.e., $100 - (100/20) = 95$) at a specified ρ . This is equivalent to the contribution of $\lambda_1 > 20 \lambda_2$.

In analyzing the factors that control population dynamics, it is interesting to know how a change in any of the survival or fecundity elements, a_{ij} , of a Leslie matrix affects the long-term growth rate given by the dominant eigenvalue. Since each a_{ij} is of a different magnitude from the others, a 10% change in one does not have the same effect on λ as a 10% change in another. To isolate the proportional change in λ resulting from a change in a_{ij} we need to examine the partial derivative of λ with respect to a_{ij} . The proportional change in λ resulting from a change in a_{ij} is known as elasticity (Caswell 2000) and is defined as:

$$e_{ij} = \frac{a_{ij}}{\lambda} \frac{\partial \lambda}{\partial a_{ij}} \quad (4.9)$$

Fig. 4-2 Elasticity is the rate of change of the dominant eigenvalue with respect to an element of the Leslie matrix.



The values of a_{ij} are, of course, specified empirically through measurements of fecundity and survival rates and λ is found in the usual way after assembling the Leslie matrix, but how do we find $\partial\lambda/\partial a_{ij}$? One way is to vary a_{ij} by a certain percentage and recalculate λ . It is probably best to do this several times. The slope of a plot of λ against a_{ij} (or the slope of a regression of λ against a_{ij}) then gives an approximation to $\partial\lambda/\partial a_{ij}$ (Fig. 4-2).

But there is a more precise way. Recall from Chapter 2 that matrices have not only right eigenvectors, but also left eigenvectors, which is a row vector by which the matrix is multiplied on the left. A left eigenvector is the right dominant eigenvector of the transpose of the matrix (to transpose a matrix make each row become the corresponding column; e.g., row 1 becomes column 1, and so forth). Call the left eigenvector of L v and the right eigenvector w . Then we can calculate $\partial\lambda/\partial a_{ij}$ as:

$$\frac{\partial\lambda}{\partial a_{ij}} = \frac{v_i w_j}{v \cdot w}$$

where $v \cdot w$ is the dot product of v and w :

$$v \cdot w = v_1 w_1 + v_2 w_2 + \cdots + v_k w_k$$

(Ellner and Guckenheimer 2006).

The left eigenvector also expresses the reproductive value of each age class (Caswell 2000). The reproductive value is the reproductive contribution of each age class to future generations, and depends on the age class's survival probability, the probability that it gives birth, and the growth of the population. To see why the left eigenvector assembles the reproductive values of each age class, assume that each age class has a reproductive value, v_i . Assume also that the young-of-the-year do not give birth, but all older age classes do. Therefore, the reproductive value of the young-of-the-year, v_1 , is the probability that it will survive to age two, s_1 , times the reproductive value of age 2, v_2 . But since the population has also grown at a rate λ , we must divide by the growth rate to discount it. So:

$$v_1 = \frac{s_1 v_2}{\lambda}$$

The reproductive value of each older age class is the reproductive value of the young-of-the-year that they give birth to times the birth rate, b_i , plus the probability that it

will survive to the next age class, s_i , times the reproductive value of that age class, v_{i+1} , again discounted by the population growth rate λ . So for the older reproductive classes:

$$v_i = \frac{b_i v_i}{\lambda} + \frac{s_i v_{i+1}}{\lambda}$$

For k age classes, this can be written in matrix form:

$$[v_1 \ v_2 \ v_3 \ \cdots \ v_{k-1} \ v_k] = [v_1 \ v_2 \ v_3 \ \cdots \ v_{k-1} \ v_k] \begin{bmatrix} 0 & b_2 & b_3 & \cdots & b_{k-1} & b_k \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & s_{k-1} & 0 \end{bmatrix} \frac{1}{\lambda}$$

Multiplying both sides by λ , we get

$$\lambda [v_1 \ v_2 \ v_3 \ \cdots \ v_{k-1} \ v_k] = [v_1 \ v_2 \ v_3 \ \cdots \ v_{k-1} \ v_k] \begin{bmatrix} 0 & b_2 & b_3 & \cdots & b_{k-1} & b_k \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & s_{k-1} & 0 \end{bmatrix}$$

which meets the definition of a left eigenvector. If the left eigenvector is normalized to sum to 1, then each element v_i gives the relative reproductive contribution of each age class to future generations.

Exercise 4.2

A biennial plant does not set seed the first year, has a 25% chance of surviving the first year, and all survivors set one seed the second year, then die.

4.2a Assemble the Leslie matrix, calculate the eigenvalues, and determine how each eigenvalue contributes to population dynamics.

4.2b Calculate the fecundity rate for the second year to exactly balance the 25% chance of surviving the first year and therefore for the population size to remain steady? (hint: what eigenvalue is needed for population size to remain steady?)

4.2c From this example, can you formulate a general conjecture about the structure of a 2×2 Leslie matrix that guarantees a stable population size? Test or prove this conjecture or reject it. What about for a 3×3 Leslie matrix? (Hint: when the population size remains the same, then $\mathbf{n}_t = \mathbf{n}_{t+1}$ and each of the elements in $\mathbf{n}_t$ equals the corresponding element in $\mathbf{n}_{t+1}$.) Do you see a pattern developing?

4.2d From the patterns in 4.2c, develop a general formula for the conditions on survival and birth rates for a Leslie matrix of any dimension that separate geometric growth ($\lambda > 1$) from geometric decline ($\lambda < 1$). These conditions give the critical combination of birth and survival rates that yield a bifurcation between extinction and unbounded growth. Now show that a 4×4 matrix also meets the conditions specified by the formula, thereby proving by induction your conditions for all matrices of any number of age classes (i.e., if you prove that the formula holds for a 3×3 matrix and then increase the size of the matrix by 1 dimension and show that the formula still holds, then it will hold for the next size matrix in this series, and the next, ...).

Life-stage-structured models

Leslie matrices apply best to species that reproduce in distinct age classes, for example vertebrates such as ourselves and many of the animal species we eat. It does not apply well to species that have complex life cycles between different stages, some reproductive and some not, such as the life cycles of mosses, liverworts, ferns, many species of higher plants, slime molds, insects, and many other “lower” forms of life which still form the majority of species on Earth.

Transitions between life stages in species with complex life cycles do not advance through a sequence of transitions from the first year to the last, as depicted in a Leslie matrix. Instead, these species have complicated transitions in which any stage can potentially change or give rise to any other stage, depending on genetic mechanisms that switch growth between different stages with the onset of environmental signals. Bonner (1965) is an excellent set of essays on the biological problems of species with complex life cycles.

Clearly, a different type of transition matrix is required to help us think about species with complex life cycles so different than our own. A related but different form of matrix model, known as a life-stage model (Caswell 2000) or a Lefkovich matrix (Lefkovich 1965), is a useful mathematical tool for thinking about problems of species with complex life cycles.

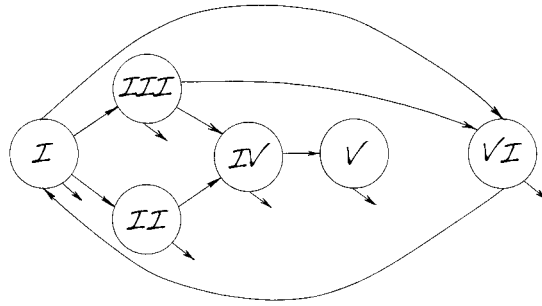
A life-stage matrix model includes a vector, not of the distribution of a population amongst age classes as in a Leslie model, but of its distribution amongst life stages. This vector is projected forward by a matrix of transition probabilities between stages. In a life-stage matrix, nonexistent transitions between two stages having a value of 0 and all other allowable transitions have values in the range (0 1.0]. The matrix has as many rows and columns as there are life stages.

A life-stage matrix has a distinctly different structure from a Leslie matrix. Although the first row still represents the number of seeds, spores, or other forms of offspring produced by each stage in the life cycle, the other rows represent the probabilities of transitions between stage j in column j to stage i in row i . Unlike a Leslie matrix, the off-subdiagonal elements of a life-stage matrix are typically not entirely occupied by zeros. In a life-stage matrix, the diagonal represents the probability that an individual in a particular stage will remain in that stage during the next time step. This probability may be zero for short life stages whose duration is less than or equal to one time step, but is typically greater than 0 and less than 1.0. (If a diagonal element were equal to 1.0, the individuals would never move out of that life stage and a life cycle would not be possible.)

As with Leslie models, the dominant eigenvalue of a life-stage model gives the long-term population growth rate and the corresponding dominant eigenvector gives the proportional distribution of the population amongst different life stages. Usually, not all life stages have the same effect on the dominant eigenvalue (Caswell 2000). If different life stages occur in different habitats, then identifying the life stage(s) that have the greatest effect on the dominant eigenvalue by calculating the elasticity can help conservation biologists identify the habitats most in need of preservation in order to conserve the species (Griffith and Forseth 2005).

Eriksson (1988, 1992) are particularly good examples of analyses of the vegetative life history stages of the dwarf shrub *Linnaea borealis* (twinflower), a common occupant of boreal conifer forests of Minnesota, Canada, New England, Scandinavia, and Siberia. This was the favorite flower of Carl Linnaeus, and so he named it for himself, much as a painter signs a painting of which he or she is especially proud.

Fig. 4-3 Structure of the life stages of *Linnea borealis*, from Eriksson (1988). I = leaf axil buds, II = flowering shoots, III = 1-year-old photosynthetic shoots, IV = 2-year-old lateral photosynthetic shoots, V = 3+-year-old lateral photosynthetic shoots, VI = main shoots.

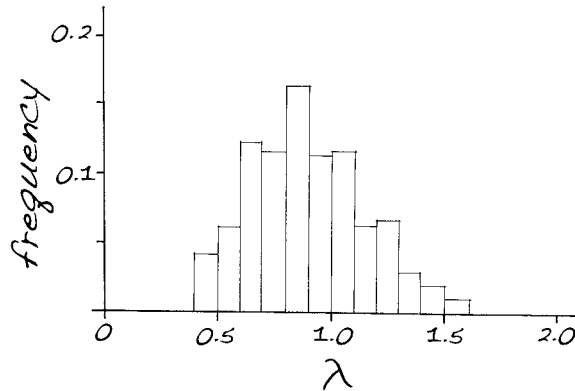


L. borealis has leaf axil buds that can develop into 1-year photosynthetic shoots; short lateral shoots; main shoots, which are long shoots trailing across the ground like vines and produce from 5 to 14 leaf axils per year; and flowering shoots which produce twin pink flowers. One-year-old photosynthetic shoots can develop into short, 2-year-old lateral photosynthetic shoots, which in turn can live for 3+ years old. The long, trailing main shoots develop only from leaf axil buds or from 1-year-old photosynthetic shoots. The short lateral shoots that are > 1 year old and the flowering shoots do not turn into the main shoots. Seeds produced by the flowering shoots have an extremely low germination rate (Eriksson 1992, Cullina 2000) and virtually all of the spread of a population is by vegetative growth of the five nonflowering shoots (Eriksson 1988). Thus, there are six life history stages, leaf axil buds and five types of shoots, all related by a web of interactions rather than a sequence of transfers as in an age-structured Leslie matrix. The flow diagram of the life history of *L. borealis* is shown in Fig. 4-3. Note the positive feedback between leaf axils (stage I) and main shoots (stage VI), sometimes via 1-year-old photosynthetic shoots (stage III) as an intermediate step. The path leading to photosynthetic shoots of 3+ years old is a negative feedback to population growth because it is a dead end that does not result in more leaf axils. This dead end and the mortality of each stage are the negative feedbacks balancing the positive feedback loop of leaf axil ↔ main shoots.

Eriksson compiled data on transitions between different life stages for a population of twinflower in Sweden by the simple expedient of marking individual shoots with colored threads corresponding to different shoot types each year. The number of threads tied on each shoot represented the number of years it had been recorded. Each two adjacent threads of a different color represented a transition from one life stage to another (and therefore data for the off-diagonal elements of rows other than row 1). Two adjacent threads of the same color represented a stage that did not change (and therefore data for the diagonal elements). Original research does not always require large grants – imagination and a trip to a hardware or dry goods store can do the trick!

A representative matrix of average transition probabilities and annual per capita production of leaf axil buds by main shoots for the population Erickson studied in Sweden looks like this:

Fig. 4-4 Frequency distribution of λ amongst different subpopulations of *Linnea borealis*, redrawn from Eriksson (1992).



From:

	I	II	III	IV	V	VI
To:	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 6.81 \\ .084 & 0 & 0 & 0 & 0 & 0 \\ .292 & 0 & 0 & 0 & 0 & 0 \\ 0 & .128 & .46 & 0 & 0 & 0 \\ 0 & 0 & 0 & .353 & .353 & 0 \\ .019 & 0 & .033 & 0 & 0 & .596 \end{bmatrix}$					

Note that the columns of rows that represent transition probabilities (i.e., other than row 1) do not add to 1 because some of the individual shoots in different stages simply die without being transformed into anything. The first row, in which the only nonzero entry is the number of new axil buds per main shoot (6.81), is analogous to the fecundity row of a Leslie matrix. Therefore, if we have a vector of numbers of different stages of twinflower in a population at time t , we can project the growth of the population forward by simple iteration of the life-stage matrix times the vector just as for Leslie matrices.

The dominant eigenvalue for the entire population was 0.836. Since this is < 1 , the population must be, on average, declining. However, by constructing different life-stage matrices from different subpopulations, Eriksson was able to estimate a frequency distribution of dominant eigenvalues (Fig. 4-4). Note that some of these are > 1 . Therefore, spatial variation in the dominant eigenvalue could yield a persistent population even though the average growth rate is < 1 .

What does this mean for the spatial dynamics of the population? Some portions with $\lambda < 1$ will obviously decline until they are locally extinct, but others with $\lambda > 1$ will expand. Therefore, the patches of *L. borealis* will creep across the forest floor and holes may even appear and refill as different subpopulations die or expand.

Summary: what have we learned?

The Leslie matrix and stage-structured matrix models suffer from the same problem as geometric growth models because they are essentially the sum of k geometric growth models, one for each of k age classes or life stages. Relaxing the assumption of constant per capita birth and death rates and making them differ among classes (but

remain constant with time within that class) does not solve the problem of the instability of the geometric model when net growth is positive. Although the growth of the population continues unbounded, the proportional distribution of the population amongst different age classes or life stages does approach a stable distribution which is given by the dominant eigenvector.

Once ecologists became familiar with matrix algebra and Leslie's paper in the 1970s, there was a flurry of activity in constructing and analyzing matrix models (see Caswell 2000 for a comprehensive review). However, once it was also realized that such models still had unstable population growth (although a stable age or stage distribution), research into them slacked off a bit. Recently there has been a renaissance of their use with regard to rare species (Griffith and Forseth 2005) and invasive species (Parker 2000). It is a straightforward experimental procedure to get fecundity and survival parameters for such species and short-term projections of their populations are of most interest. Although per capita birth and survivorship rates do not remain constant forever, we can assume that they remain reasonably constant for some short sequence of time intervals over which policy needs to be made to save a local population from going extinct. The projections of Leslie or life-stage matrices can then be used to make short-term forecasts to answer some important questions in conservation biology: Will the rare species population go extinct in the near future? Will the invasive species population explode in the next few years? Is the eventual age structure of a population (especially a harvested one) desirable – the dominant eigenvector may or may not represent a desirable state of the natural population. Examination of eigenvalues can also reveal unsuspected oscillations in age class distribution and total population density if they are complex conjugates, thus helping us avoid surprise when age classes change unexpectedly. Finally, the elasticity of the dominant eigenvalue to particular transitions can also reveal especially important age classes or life stages upon which conservation efforts should focus. Leslie and life-stage matrices are appropriate for when short-term projections and quick action is needed or, in Caswell's (2000) phrase, for forecasting the consequences of a policy rather than for accurate prediction. Perhaps as more case studies of applications of Leslie and life-stage matrices to rapidly growing or declining populations are assembled, we will begin to see some patterns in their properties that characterize these species, on which there is now much focus in conservation biology and natural resource management.

Leslie and life-stage matrices can also be used to design experiments to test how different factors that are hypothesized to affect birth and survivorship of a given age class in turn control overall population dynamics – for example, how different angling pressures or size limits on release of fish could increase, decrease, or cause oscillations in a game fish population. Another example could be to determine how different harvesting practices in a woodland where the decision whether to cut a tree or not depends on the tree's age or size class and affects the dynamics over the next several years. The effect of these experimental manipulations in the short run can be tested by short range projections of an initial vector of age class distribution by the appropriate matrix. Although the dynamics of natural populations will almost never conform to the dominant eigenvalue and eigenvector (because these, too, will change if birth and survivorship rates are not constant), nonetheless the dominant eigenvalue and eigenvector are good relative indices that can be used to compare population dynamics of the same species but in different environments, thus

determining how specific environmental conditions determine population growth rates and age class distribution. Examples of empirical studies using matrix population models can be found in Caswell (2000).

Open questions and loose ends

The problem with matrix population models, the same problem which plagued exponential and geometric models of homogeneous populations, is that per capita birth and survival rates are constant. Only in the very special case of a matrix with a dominant eigenvalue equal to 1 is the population stable. DeAngelis et al. (1980) and Caswell (2000) suggest that relaxing the assumption of constancy of birth or survival/transition rates and making these entries functions, for example of the density of one or more age or stage classes. This is one promising avenue for further exploration of matrix models.

Matrix models can also be used to analyze two subpopulations which are coupled through immigration and emigration. Such is the case for so-called source ($\lambda > 0$) and sink ($\lambda < 0$) populations in which individuals from the source migrate to the sink. To stimulate your thinking, consider the following toy model of a source and a sink population of a biennial species. Suppose a fraction, e , of the year-2 individuals of the source population, y , emigrate each year before reproducing. This reduces the net effective per capita birth rate of y_2 by a fraction e :

$$\begin{bmatrix} 0 & b_{2,y} - e \\ s_{1,y} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

If a certain fraction μ of these emigrants find their way to the x sink population and reproduce there, then the Leslie model for x would need to be modified like this:

$$\begin{bmatrix} 0 & b_{2,x} \\ s_{1,x} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 + \mu e y_2 \end{bmatrix}$$

What are the conditions on b_i , s_i , μ and e that would allow the entire population $x + y$ to persist? What happens when you generalize this toy model into a larger Leslie matrix with more age classes with different emigration rates for each age class? What about life-stage matrices: are there certain life history configurations that are more easily prevented from having sink populations than others? How do such couplings between matrix models of subpopulations affect elasticity and other properties?

5 Nonlinear models of single populations: the continuous time logistic model

In the previous chapters, we saw that the exponential model either describes a population that grows to infinite size when $r = b - d > 0$ or declines to zero when $r < 0$. The first is clearly not possible and the second is extinction (or complete decomposition of dead biomass). According to these models, there is only one population size that could be maintained indefinitely, and that is a population size of 0! We also saw that relaxing the assumption of homogenous values of b and d for all age classes, resulting in the Leslie matrix model, also did not help. Although the Leslie matrix gave interesting insights into the proportional distribution of age classes, the total population size also increased indefinitely if total fecundity exceeded total mortality summed over all age classes. Life-stage models suffer similar problems. In all these linear models, there appears to be a delicate balance between extinction and unbounded population growth that we do not see in nature. Clearly something is wrong here.

And yet, the beginnings of the model – that changes in population size are equal to births minus deaths – seemed reasonable. Where did we go wrong? To find out, we must take a hard look at one of the hidden assumptions of the exponential model and propose a way to relax that assumption in a more realistic way. Indeed, this procedure will be a common thread throughout this book. We shall see that our starting point – the exponential model of population growth by birth and death of individuals – remains at the core of all more elaborate models but with its faults at least partly tamed by how we relax the assumptions. If we are lucky, relaxing hidden assumptions in biologically reasonable ways will lead not only to more realistic behavior of the model but also to richer and surprising behaviors that suggest new hypotheses. But it will also bring to light other hidden assumptions which we must then examine and relax in the hopes that, in the limit, we approach a mathematically rigorous model that captures the behavior of the natural world and which, as Einstein said, is as simple as needed but no simpler.

Derivation, forms, and interpretation

It seems axiomatic that changes in population densities are the result of births and deaths (neglecting for a while immigration and emigration) so that assumption in and of itself should not be the problem. The problem arises, however, because of the hidden assumption that the per capita birth and death rates, b and d , respectively, are constant for all members of a population (the exponential model) or for all members of an age class (the Leslie matrix) regardless of the density of the entire population. Now that we have brought this hidden assumption to light, it does not seem at all reasonable – the probability of giving birth and dying clearly changes in response to the current state of the population.

How can we relax this assumption of constant per capita birth and death rates in a way that would seem more realistic? It seems reasonable that as the population density increases, the probability of death should also increase, perhaps because of social stresses due to crowding, because of increased probability of cannibalism, or perhaps because there is a finite amount of material resources that can be “made” into more bodies. Similarly, it is also reasonable to expect that per capita birth rates should also decline, perhaps because less time must be spent on reproduction than on gathering the increasingly scarce supply of resources when they are divided amongst more individuals. Thus, per capita birth and death rates might depend at least partly on population density. This is termed density-dependent population regulation.

The mathematically simplest way to express this is to assume that birth rates decrease linearly and death rates increase linearly in proportion to N :

$$\begin{aligned} b &= b_0 - \beta N \\ d &= d_0 + \delta N \end{aligned} \tag{5.1}$$

where b_0 and d_0 represent some nominal per capita birth and death rates when N is very small and β and δ describe how per capita birth and death rates change as N increases. β and δ therefore describe the strength of the density-dependence of per capita birth and death rates. Now let's substitute these expressions for b and d into the differential equation that resulted in exponential decay,

$$\frac{dN}{dt} = (b - d)N$$

and see where this leads us:

$$\frac{dN}{dt} = [(b_0 - \beta N) - (d_0 + \delta N)]N$$

By gathering terms in N inside the brackets we get:

$$\frac{dN}{dt} = [(b_0 - d_0) - (\beta + \delta)N]N \tag{5.2}$$

Now, when $dN/dt = 0$ there are two equilibrium points, either $N_1^* = 0$ (the so-called “trivial” equilibrium) or else N_2^* when the expression in brackets = 0. N_1^* is the same as the equilibrium point for the exponential model, but relaxing the assumptions of constant per capita birth and death rates resulted in a new equilibrium, N_2^* . As we shall see throughout this book, relaxing assumptions often cause new equilibria (or at least highly modified equilibria) to appear. Identifying these new or modified equilibria and examining their relationship to any “old” equilibria of the unmodified model is always a productive path to follow.

So, let's find the new equilibrium point by setting the term in brackets equal to 0 and solve for N_2^* :

$$0 = (b_0 - d_0) - (\beta + \delta)N_2^*$$

$$N_2^* = \frac{b_0 - d_0}{\beta + \delta}$$

Now recognize that the numerator of the right hand side is simply the net per capita growth rate, r , when N is small:

$$r = b_0 - d_0 \quad (5.3)$$

So, let's also rename this equilibrium K (the reason for this will be obvious in a moment), yielding:

$$N_2^* = K = \frac{r}{\beta + \delta} \quad (5.4a)$$

or

$$\beta + \delta = \frac{r}{K} \quad (5.4b)$$

These are the key intermediate steps in the derivation. Most derivations and proofs have one or more key steps about which the whole procedure turns – the trick to understanding a proof or derivation is to find and understand these key steps. As we shall see, these steps, especially Eq. 5.4, have important implications for interpreting what follows.

We can now simplify Eq. 5.2 by substituting Eqs. 5.3 and 5.4, yielding

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right) \quad (5.5)$$

which you should recognize as the famous logistic growth model for populations. By means of these substitutions, we have reduced the number of parameters from four (b_0 , β , d_0 , δ in Eq. 5.2) to two (r and K in Eq. 5.5), thus simplifying the model's appearance (but also introducing some difficulties of interpretation that we shall see shortly). Note that the original exponential growth model (rN) from which we started is still part of the logistic model, only now modified by the term in parentheses which includes the new parameter (K) which is in turn a function of all four of the new parameters we needed to introduce to relax the assumption of constant per capita birth and death rates.

This derivation of the logistic model from the exponential model was first presented by Kostitzin (1937), although Verhulst* (1838) originally presented it in another form:

$$\frac{dN}{dt} = rN - aN^2$$

where $a = r/K$. This form can actually be derived by assuming that the net per capita growth rate r (the difference between per capita birth and death rate) decreases linearly with increasing population density by some proportionality constant, a . That is,

*See Hutchinson (1978) and Cramer (2003) for interesting histories of the logistic equation, especially about Verhulst and the Napoleonic Empire.

$(r - aN)$ is substituted for r in the exponential model, $dN/dt = rN$. While this is certainly mathematically legitimate, biologically it confounds the two processes of birth and death into one parameter, net survival rate r . Births and deaths are determined by different behaviors and environmental influences, and so confounding the two into one parameter makes the derivation much simpler but at the expense of obscuring the more fundamental biological processes of birth and death and the different ways that each of these processes are affected by population density. The derivation we have used here seems more fundamental to me because it stays closer to the basic processes of birth and death during the derivation.

Another common form of the logistic model reduces the number of parameters from two (r and K) to just one (r):

$$\frac{dN}{dt} = rN(1 - N) \quad (5.6)$$

It is often claimed, erroneously, that Eq. 5.6 is “derived” from Eq. 5.5 simply by “setting $K = 1$ ”. While this has the anticipated effect in this case, this procedure is mathematically cavalier and can often lead to mathematical problems and problems of interpretation in many other cases where a parameter is “set to 1” or “set to 0” (Gurney and Nisbet 1998). One proper way to reduce the number of parameters in a model (and thus aid in its mathematical analysis) is to redefine one or more of the variables to include one or more of the parameters such that the parameters are removed when the redefined variable is substituted into the original equation. For example, if we define a new variable $N' = N/K$ (or $N = KN'$) and then substitute KN' for N (or N' for N/K) in Eq. 5.5:

$$\frac{d(KN')}{dt} = rKN'(1 - N')$$

and using the chain rule for the derivative of a product to factor out K on the r.h.s:

$$K \frac{dN'}{dt} = rKN'(1 - N')$$

then by canceling K on both sides we arrive at an equation which is formally equivalent to Eq. 5.6 but where the units of N' (population density as a fraction of K) are not the same as those of N (population density in numbers of individuals).

In whatever form, the logistic model has played an important role in theoretical ecology,* perhaps more important than any other equation or set of equations with

*An interesting use of the logistic equation has recently been made by Cisne (2005a), who used it to estimate survival of Medieval scientific and mathematical manuscripts written by The Venerable Bede (see also Gilman and Glaze 2005, Declerq 2005, and Cisne 2005b). The copying of manuscripts in monasteries was taken as the process of birth and the destruction of manuscripts was taken as the process of death. Similarly, Gingerich (2005) presents a fascinating study of documenting the survival and provenance of original copies of Copernicus's *De Revolutionibus* which uses an exponential “decay” model to back-calculate how many copies were originally printed.

the possible exception of the Lotka–Volterra equations (Kingsland 1995). Pearl (1925) elevated it to the status of a biological “law.” However, ecologists have had a stormy relationship with the logistic model (Olson 1992), some (Pearl 1925, Feller 1940, Hutchinson 1978) being entranced with its elegance and ability to provide a very good fit to time sequences of growth of isolated populations while others (e.g., Ginzburg 1992, Getz 1996) decrying its unrealistic assumption regarding how per capita birth and death rates depend linearly on population density N . Much of this controversy can be put into perspective if one keeps in mind how the logistic model was derived from the exponential model, especially the assumption of linear dependence of per capita birth and death rates on population density and also the relationship between r and K embodied in Eq. 5.4a. But before we delve into the implications of Eq. 5.4a and a more detailed examination of the behavior of the logistic model, perhaps you would like to try your hand at relaxing the assumption of density-dependence of b and d in a different way.

Exercise 5.1

We derived the logistic model by assuming that b and d of the exponential growth model vary linearly with population density, N . But there is no reason to think that this is the only way that the assumption of constant per capita birth and death rates could be relaxed. One could assume some more complicated, but still reasonable, relationships between per capita birth and death rates and population density. For example, the birth and death rates might change in proportion to the frequency of any two individuals meeting and interfering with each other in finding mates or food. For example, cannibalism would increase adult per capita death rates or decrease per capita birth rates if very young individuals were consumed before they reach reproductive age. Preempting nest sites would also effectively decrease per capita birth rates (some spiders actually do all these things – see Wise 1993). Each of N individuals can potentially meet and interfere with $N - 1$ individuals and so the total number of interfering pairs is $N(N - 1)$, or, if N is large, approximately N^2 . Following the procedure above, derive a new “logistic-like” equation assuming that per capita birth rates decline and per capita death rates increase linearly with N^2 , respectively, rather than simply with N :

$$b = b_0 - \beta N^2$$

$$d = d_0 + \delta N^2$$

How does this new model differ from the standard logistic model? How does K in this new model compare with K in the standard logistic model? Can you think of other ways to relax the assumption of constant per capita birth and death rates?

Thus far, we have (rather obviously) avoided calling K “environmental carrying capacity” or simply “carrying capacity.” The reason is that we made no explicit assumptions about the “environment” in deriving the logistic model except that the population is its own and sole environment! Indeed, we have not even specified a mechanism by which the population or environment alters per capita birth and death rates – we just postulate a functional form of the effect of N on each (and, if you did Exercise 5.1, you will have seen that other plausible functional forms yield different, but perhaps equally valid, logistic-like models). Strictly and mathematically speaking, K is only the nontrivial equilibrium of Eq. 5.2. It appears in the standard form of the logistic model (Eq. 5.5) as a parameter substitution that helps clean up its appearance. Of course, it is possible that K might be related in some way to a specific environmental property (see Hutchinson 1978 for some examples), but that it is a hypothesis to be tested by experiment. It is not a conclusion that necessarily follows from the mathematics. The terms “environmental carrying capacity” and “carrying

capacity” have become so embedded in ecology (and especially in the practice of wildlife management) and so tightly linked to the logistic model that it is well to remember that strictly speaking K has nothing necessarily to do with the environment outside the population of interest!

One way to make some environmental variable explicit in the logistic model is to develop an empirical relationship between K and some measurement of the environment, and then substitute this empirical function for K in the logistic model (see Hutchinson 1978 for examples). However, it would be mathematically more correct, and biologically more insightful, to develop specific hypotheses about how some environmental variable might affect each or all of the underlying “hidden” parameters, b_0 , β , d_0 , and δ , that specifically describe the underlying biological processes of birth and death rates and their density-dependence, then insert such a function into the appropriate place in Eq. 5.2 and proceed with the derivation as above. The new non-trivial equilibrium will then naturally include the hypothesized effect of the environment on the appropriate underlying biological processes of birth rates, death rates, and their density-dependence.

The relationship between r and K in Eq. 5.4a, a key intermediate step in the derivation, also has important implications for interpreting the logistic model. This equation specifies that r and K are proportional to one another, the proportionality constant being the reciprocal of $(\beta + \delta)$. However, it is often assumed that r and K are independent of one another. In fact, it is r and $(\beta + \delta)$ which are independent, not r and K . If the logistic model were viewed simply as a quadratic function without biological meaning, then r and K are certainly two parameters that could be treated independently of one another. But the derivation of the logistic model from the assumption of density-dependent per capita birth and death rates, during which Eq. 5.4a is the key step, says otherwise. If you assume that β and δ are constant and independent parameters (as we did in Eq. 5.1) and then derive the logistic model, then r and K are correlated (Eq. 5.4a). If you wish to determine how changes in r affect the behavior of the logistic model while holding K constant, then Eq. 5.4a requires that you also need to adjust the density-dependent terms, $\beta + \delta$, to compensate and maintain the constancy of K . But then we are necessarily altering the constancy of the density-dependent terms, which is the assumption we began with. The derivation invests the model with biological meaning, but we pay a price for that biological meaning in Eq. 5.4a – in effect, we lose a degree of freedom (independence of r and K) by investing the logistic model with the biological meaning of underlying linear density-dependence of birth and death rates. You have to choose one or the other assumption. There is no free lunch!

Gabriel et al. (2005) call the interpretation of the logistic model in which r and K are correlated the “constrained” logistic model and the interpretation in which they are independent the “free” logistic model. This would be merely a matter of semantics except that they also show how the constrained interpretation avoids some criticisms of the logistic model raised by Levins (cf. Hutchinson 1978) and Ginzburg (1992), who each assumed the free interpretation.

Explicit solution of the logistic model

The logistic model is the simplest nonlinear differential equation, especially in the one parameter form (Eq. 5.6). Most nonlinear differential equations cannot be explicitly solved in a general time-dependent form – that is, there are often no methods to integrate a differential equation for dN/dt to obtain an explicit equation for $N(t)$ in terms of initial conditions, $N(0)$, and the parameters – or else the explicit, time-dependent

solutions are too unwieldy to be of much help. But the logistic model has an explicit, time-dependent solution and it is worth finding it not only to demonstrate the power of some of the algebra and calculus covered in Ch. 2, but also to demonstrate its relation to the exponential growth model in more detail.

Using the form of Eq. 5.5,

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right)$$

we separate the terms in N on the r.h.s and those in t on the l.h.s:

$$\frac{dN}{N(1 - N/K)} = rdt$$

then integrate the r.h.s from $N(0)$ to $N(T)$ and the l.h.s. from $t = 0$ to $t = T$:

$$\int_{N(0)}^{N(T)} \frac{dN}{N(1 - N/K)} = \int_0^T rdt$$

The integral of the r.h.s. is simple after factoring out the constant r , but we must integrate the l.h.s. by dividing the expression into the sum of two parts, then integrate each part separately:

$$\int_{N(0)}^{N(T)} \frac{1}{N} dN + \int_{N(0)}^{N(T)} \frac{1/K}{1 - N/K} dN = r \int_0^T dt$$

Recall that the integral of $1/N$ is $\ln N$ and from a table of integrals we find that the integral of the second term on the l.h.s. is $-\ln(1 - N/K)$. To find definite integrals we evaluate at the upper and lower bounds and subtract the expression for the lower bound from that of the upper:

$$\ln(N(T)) - \ln(N(0)) - \ln(1 - N(T)/K) + \ln(1 - N(0)/K) = rT$$

Now we get rid of the $\ln$ expressions by exponentiating both sides, recalling from Chapter 2 that $e^{a+b} = e^a e^b$, $e^{-a-b} = 1/e^{a+b}$, and $e^{\ln(a)} = a$:

$$\frac{N(T)(1 - N(0)/K)}{N(0)(1 - N(T)/K)} = e^{rT}$$

Solving algebraically for $N(T)$ and assuming that $N(0)$ represents the initial conditions, we get:

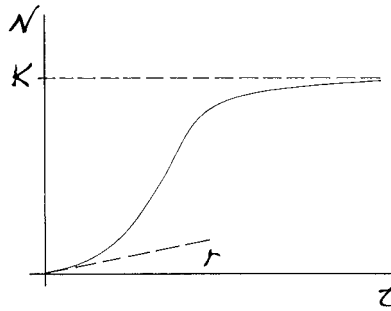
$$N(T) = \frac{N(0)e^{rT}}{1 + \frac{N(0)(e^{rT} - 1)}{K}} \quad (5.7a)$$

or in an alternative form:

$$N(T) = \frac{K}{1 + \left(\frac{K - N(0)}{N(0)} \right) e^{-rT}} \quad (5.7b)$$

which yields the familiar “S-shaped” curve (Fig. 5-1).

Fig. 5.1 The growth of a population according to the logistic model.



Note that the solution for the original exponential model remains as part of the solution for the logistic model in the numerator of Eq. 5.7a. The denominator modifies the solution of the exponential model according to the new parameters that follow from relaxing the assumption of constant b and d in Eq. 5.1 (try inserting the expressions for r and K in terms of b_0 , β , d_0 , and δ from Eqs. 5.3 and 5.4a into Eq. 5.7 to more clearly understand this).

Let's see how this solution behaves at $T = 0$ and as $T \rightarrow \infty$. Assume that at $T = 0$, $N(0) = 0$. Then clearly both $N(T)$ from Eq. 5.7a and dN/dt from Eq. 5.5 equal 0. Therefore, $N(0) = 0$ is an equilibrium. But if we introduce a single reproductive individual such that $N = 1$ and let $T \rightarrow \infty$, from Eq. 5.7b we can see that $e^{-rT} \rightarrow 0$ causing the denominator to also go to 1 and $N(T) \rightarrow K$, the other equilibrium solution. Therefore, if we introduce at least one reproductive individual into the population the solution moves from one equilibrium (0) to the other (K) as $T \rightarrow \infty$. We thus come now to the question of the stability of the logistic model after a perturbation such as introducing (or harvesting) at least one individual into (or from) the population.

Geometric analysis of the stability of the logistic model

To begin a geometric stability analysis, we graph dN/dt against N , thus showing how the rate of change of the population varies with population density, N . Since the logistic model is a quadratic equation, the graph is simply a parabola that crosses the N -axis at the two equilibrium points, 0 and K , where $dN/dt = 0$ (Fig. 5-2).

We now place arrows to the right or left to each side of each equilibrium point in the direction of population change specified by the sign of the derivative (Fig. 5-3). Notice that arrows point away from 0 on either side but they both point towards K from either side. Either introducing (stocking) at least one individual or harvesting one individual in the vicinity of K will always result in the population returning to K . Doing the same around 0 will always result in the population moving away

Fig. 5-2 dN/dt of the logistic model changes parabolically with increased N . Note that 0 and K are the equilibria when $dN/dt = 0$.

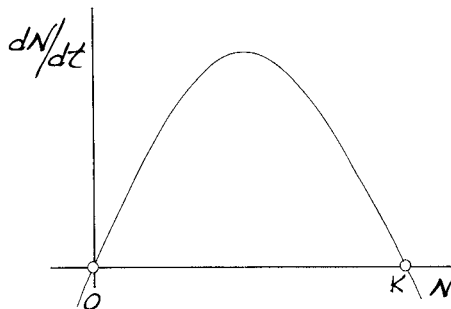
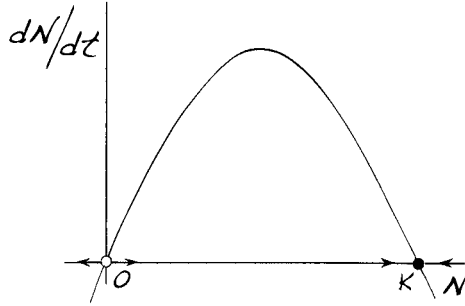


Fig. 5-3 K is a stable equilibrium and 0 is an unstable equilibrium.



from 0 , so long as $r > 0$ (a point to which we shall return shortly). Therefore, K is a stable equilibrium point while 0 is an unstable equilibrium point.

Algebraic analysis of the stability of the logistic model

Let's now calculate the eigenvalues of the logistic model at 0 and K and see if they shed more light on the geometric analysis of their stabilities. To calculate the eigenvalues at each equilibrium, we take the derivative of the logistic model with respect to N and evaluate it at the two equilibrium points, 0 and K . The derivative of the logistic model with respect to N is:

$$\lambda = \frac{df(N)}{dN} = r \left(1 - \frac{2N}{K} \right) \quad (5.8)$$

(verify this for yourself). Unlike the eigenvalue of the exponential model, which only depended on the parameter r , both the sign and magnitude of this eigenvalue depend on N , and so must be evaluated at particular values of N . To evaluate this for both equilibria, we insert 0 or K into the r.h.s. of Eq. 5.8 and solve, yielding:

$$\left. \frac{df(N)}{dN} \right|_{N^*=0} = \lambda_0 = r \left(1 - \frac{0}{K} \right) = r$$

and

$$\left. \frac{df(N)}{dN} \right|_{N^*=K} = \lambda_K = r \left(1 - \frac{2K}{K} \right) = -r$$

The eigenvalues r or $-r$ are the values of the slopes of the lines tangent to the logistic function at 0 and K , respectively. Unlike the geometric analysis above, which only told us whether the equilibria were stable or unstable, the calculations of the eigenvalues near these two equilibria tell us quantitatively what controls the fate of perturbations near equilibria, namely r .

Recall that the equilibrium for the exponential model was also 0 . The eigenvalue for the exponential model was also simply r but it was independent of the value of N . Therefore, the exponential equation was globally unstable (unstable for all values of N) when $r > 0$ and globally stable (stable for all values of N) when $r < 0$. In the logistic model, when $r > 0$, K is stable only for $N > 0$. To see this, examine more closely the above graphs: any value of $N > K$ has a corresponding $dN/dt < 0$, but any value $0 < N < K$ has a corresponding $dN/dt > 0$. Therefore, the domain of N with which $N \rightarrow K$ is for all $N > 0$ (again assuming $r > 0$). If $N < 0$, then it decreases in an unbounded

manner and does not converge on K . The domain $N > 0$ is called the *basin* (or *domain of attraction*) for $N^* = K$. Although this is a very large domain, nonetheless it is not global (i.e., it excludes all negative numbers). Therefore, the stability of the logistic model near K is called *local stability*, the “local” being the domain of attraction $N > 0$.

Local stability is characteristic of equilibria of nonlinear models. The reason for this is that the eigenvalues of a nonlinear model are also functions of the argument of the model, in this case N , because nonlinear models include higher order terms of N . Therefore, when you take the derivative of a nonlinear function with respect to its argument, N , you also get terms in N . Eigenvalues of linear models, which do not contain higher order terms, are solely functions of the parameters and do not depend on the argument. Unlike the exponential model, here the eigenvalue equals r or $-r$ precisely only at 0 or K and approximately r or $-r$ only near 0 or K , respectively. Therefore, r or $-r$ describe stabilities only in the locales of 0 or K , or local stabilities, not global stabilities as in the linear differential equation for exponential growth. Eigenvalues far from 0 or K have other expressions that are some fraction or multiple of r (insert other values of N , such as $\frac{1}{2}K$ or $2K$ into Eq. 5.8 to see this).

The local nature of eigenvalues is an important feature of nonlinear differential equations. One must be careful to keep this in mind when speaking of the stabilities of nonlinear systems. Determining the behavior of large perturbations far from equilibria of nonlinear models often requires evaluating the derivatives of the higher order terms of the Taylor Series expansion of the model, not simply the first (linear) term as we have done here. One cannot say anything with certainty about the stability of points far from equilibria in nonlinear models using the sort of linearization employed above. Because populations and ecosystems are rarely near their equilibria, this has led some to conclude that stability analysis near equilibria is not very useful for understanding natural systems. This is perhaps too strong a conclusion. On the one hand, it is true that we cannot know for certain the details of trajectories of systems far from equilibria without more complicated analysis involving higher order terms of the equations. On the other hand, we know that systems are moving away from one or another of their unstable equilibria and towards one or another of their stable equilibria, so the eigenvalues near the equilibria define the general directions of movement of the system and the factors that control the rates of these movements.

At what value of N is the growth rate maximum and what is the growth rate at this population density? To find this, we set the derivative with respect to N (Eq. 5.8) equal to 0 and solve for N . It is an easy matter of some algebra to see that the maximum value of dN/dt happens when $N = \frac{1}{2}K$ (verify this for yourself by setting Eq. 5.8 to 0 and solving for N). Substituting $\frac{1}{2}K$ for N in the logistic model (Eq. 5.5) we find that the maximal growth rate of the population at $\frac{1}{2}K$ equals $\frac{1}{4}r/K$.

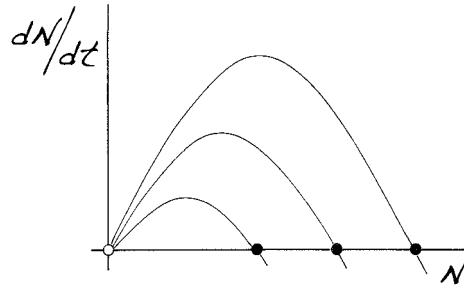
Exercise 5.2

Analyze the stability of the logistic-like model you derived in Exercise 5.1 using the above approach of geometric analysis and linear stability analysis by means of eigenvalues evaluated at the equilibria.

Behavior of the logistic model at critical values of r : the transcritical bifurcation

Having seen that r plays such a dominant role in the time-dependent behavior and stabilities of the equilibria of the logistic model, we next ask how does the behavior of the logistic model change with different values of this parameter? Since K is positively correlated with r , whatever conclusions we come to for r also apply to K . However, because r is more simply expressed in terms of the fundamental

Fig. 5-4 Behavior of the logistic model as r and K simultaneously decrease.



parameters b_0 and d_0 (Eq. 5.3) whereas K is a more complicated function of b_0 , β , d_0 , and δ (Eq. 5.4a), the analysis is simpler to understand if we choose r as the parameter that we control.

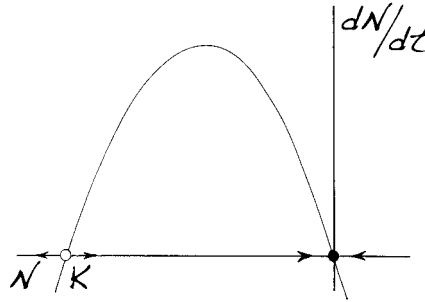
For the exponential model, we have seen that as r passed through 0 from positive to negative, the equilibrium point ($N^* = 0$) switched from being unstable to being stable. $r = 0$ was therefore a critical value in the exponential model around which there was a sudden change in the qualitative behavior of the model and the stability of its equilibrium, which we called a bifurcation. Since we derived the logistic model from the exponential model and since 0 is also an equilibrium point of the logistic model, would the same thing happen here? What happens to the stability of K , the new equilibrium point?

Having r pass through 0 for the logistic model of population dynamics might at first seem like a mathematical exercise devoid of biological meaning, but it is not. Sink populations can be modeled as having negative r in which death rates exceed birth rates and, as we shall see, some models of mutualism use $r < 0$ to describe obligate mutualists whose populations cannot grow on their own but require the presence of a symbiont for births to exceed deaths.

We can begin by using the geometric approach to get a qualitative feel for what happens as r passes through 0. Because r and K are correlated, as r decreases, so does K (Fig. 5-4). Consequently the two equilibria approach each other until they coincide when $r = 0$. What happens as r passes through 0 and becomes negative? Since r and K must have the same sign, K also becomes negative and the graph of dN/dt vs. N moves out of the first quadrant into the second quadrant, while remaining in the third and fourth quadrants but now for different legs of the parabola. Finally, we add the arrows along the N -axis to indicate the direction of change of N according to the sign of dN/dt . There are still two equilibria as before (0 and K) only now they have exchanged stabilities, 0 now being stable and K being unstable (Fig. 5.5). This is a *transcritical bifurcation*: a bifurcation resulting in two equilibria exchanging stabilities as a parameter (in this case, r) or combination of parameters passes through some critical value. Quadratic functions such as the logistic model are the canonical or normal forms for transcritical bifurcations because they are the prototypical or simplest functions having such behavior (see Guckenheimer and Holmes (1983) and Strogatz (1994) and for more formal uses of the concept of normal forms).

We can also use the algebraic solutions of the eigenvalues evaluated at the equilibria to confirm our geometric analysis of the appearance of a transcritical bifurcation, perhaps in this case even more simply. Recall that $\lambda_0 = r$ and $\lambda_K = -r$ and so when $r > 0$, K is stable and 0 is not. But inserting values of $r < 0$ into the expressions for the

Fig. 5-5 When $r < 0$, K becomes unstable and 0 becomes stable.



eigenvalues results in $\lambda_0 < 0$, hence the equilibrium 0 becoming stable, and $\lambda_K > 0$, hence the equilibrium K (which is now also negative) becoming unstable.

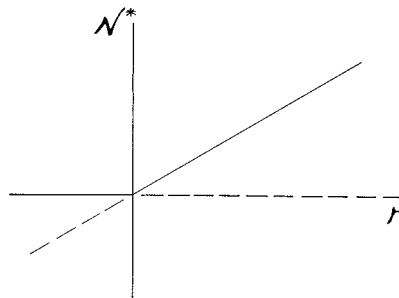
Finally, we can plot how the equilibrium solutions change with changes in the control parameter in a bifurcation diagram. For the logistic equation we have two solutions, $N_1^* = 0$ and $N_2^* = K$. The first is stable when $r < 0$ and the second is stable when $r > 0$. In a bifurcation diagram, unstable solutions are plotted as dashed lines and stable solutions as solid lines. Plotting both solutions against r yields two lines (Fig. 5-6). The line describing $N_1^* = 0$ is simply the x -axis and is solid when $r < 0$ and dashed when $r > 0$. The other line, $N_2^* = K = (1/\beta + \delta)r$ is a straight line with slope $(1/\beta + \delta)$, from Eq. 5.4a, and is dashed when $r < 0$ and solid when $r > 0$. This bifurcation diagram shows clearly that when $r = 0$ the solutions collide and switch stabilities, each line being a branch from $N_1^* = N_2^* = 0$. This is precisely how we defined a bifurcation in Chapter 3: A model is said to bifurcate from a point, usually an equilibrium point, if there are two or more solutions that approach that point as some parameter, a_1 , or some combination of parameters $g(a_1, a_2, \dots, a_k)$ approaches a critical value.

Exercise 5.3

What would a bifurcation diagram for the logistic model look like using K as the control parameter?

We shall see transcritical bifurcations in other ecological models, some which have the logistic model as their base and some which do not. We will use both geometric and algebraic analysis to uncover the critical values of parameters that separate different equilibria and lead to the appearance of new equilibria. But the above analyses for the logistic model are our first, and in many ways simplest, nontrivial example of a bifurcation analysis of a nonlinear model. Therefore, it is worth making sure you understand what we have just done.

Fig. 5-6 Transcritical bifurcation of the logistic model as r increases. The equilibrium which is stable at particular values of r is the solid line, and the unstable equilibrium is the dashed line.



Exercise 5.4

Perform a bifurcation analysis of the model you derived in Exercise 5.1 using the above geometric and algebraic techniques.

Summary: what have we learned?

Relaxing the assumptions of constant per capita birth and death rates in the exponential model results in the logistic or logistic-like models, depending on how the assumptions are relaxed. Both the differential form of the logistic model and its time-dependent solution retain the exponential model and its solution, respectively. But the logistic model also includes new terms that modify the exponential model and tame some of its more unrealistic behaviors by preventing growth from accelerating to infinite population density.

The logistic model has a richer and more realistic behavior compared with the exponential model. When $r > 0$ there is the appearance of a new equilibrium population size, K , lacking in the exponential model and which, more importantly, is also stable. The logistic model also retains the equilibrium of the exponential model, 0, which also remains unstable. With time, any population less than K moves away from 0 and towards K and any population larger than K decreases towards K in the limit as $t \rightarrow \infty$.

However, this new and richer behavior comes at a price: the parameters of the logistic model, r and K , are not independent and care must be taken in their interpretation and manipulation. Although K is usually interpreted in terms of environmental factors that affect birth and death rates, and hence is often called environmental carrying capacity, this interpretation is a hypothesis that needs experimental testing and is not intrinsic to the mathematics of the derivation. It may or may not be true depending on the particular population and the circumstances.

The parameter r controls the stability at and near the two equilibrium points. Furthermore, as r passes through 0, the two solutions switch stability, and 0 becomes stable while K , which is now also less than 0, becomes unstable. This is the simplest version of a transcritical bifurcation, which we shall see repeatedly throughout the remainder of this book.

In the past decade, there has been a resurgence of interest in generalizing the logistic model from other assumptions of density-dependence and population structure. In Exercise 5.1, you effectively derived a specific realization of a generalized logistic model proposed by Richards (1959):

$$\frac{dN}{dt} = rN \left[1 - \left(\frac{N}{K} \right)^b \right]$$

Open questions and loose ends

where the parameter b distorts the symmetric parabolic shape of the logistic model into other forms. Unfortunately, the parameter b as originally proposed has no biological interpretation, and so is simply an empirical tuning parameter that controls the shape of the function (Zeide 1993). Birch (1999) derives and analyzes a different generalized logistic model that is in many ways empirically better than the Richards model, but again by introducing a shape parameter whose relation to specific biological mechanisms is not yet clear. But, in fact as Exercise 5.1 shows, specific values of b can be derived from biological hypotheses on alternative density-dependent

mechanisms affecting birth and death rates. Ackleh et al. (1999) derive yet another generalized logistic model by assuming that birth and death rates are distributed among subpopulations according to some distribution function. They then proceed to show how natural selection operating on these distribution functions causes the entire population and its equilibrium to converge on that of the subpopulation with the highest ratio of birth to mortality. It would be interesting to see a systematic study of the derivations, behaviors, and meanings of a family of generalized logistic models and how they depend on assumptions of how birth and death rates depend on population density and population structure.

6 Discrete logistic growth, oscillations, and chaos

In Chapter 2, we saw that the behaviours of the discrete geometric model and the continuous exponential model are not precisely the same. Although the exponential model is the limit of the geometric model as $\Delta t \rightarrow 0$, the parameters that stand for net per capita growth rate in the two models, r and R respectively, are not equal and, while both models undergo a bifurcation from unstable unbounded growth to stable decay at critical values of r or R , the critical values are not the same. This suggests caution in interpreting behaviors of discrete and continuous versions of the same model, especially in assuming that they are equivalent in all their properties and behaviors. If the behaviors of continuous and discrete versions of simple linear models differ in some respects, then what about the behaviors of continuous and discrete versions of more complicated nonlinear dynamic models? Here, we will see that their behaviors can be radically different.

Let's first derive several versions of a discrete model of population growth that are counterparts to the continuous logistic model we derived in Chapter 5.

The discrete quadratic model

Recall that the geometric model of growth was:

$$N_{t+1} = N_t + BN_t - DN_t = (1 + B - D)N_t$$

As before, assume birth rate decreases linearly with population size and death rate increases linearly with population size and substitute these expressions into the mass balance equation for population growth:

$$N_{t+1} = N_t + (B_0 - bN_t)N_t - (D_0 + dN_t)N_t = [(1 + B_0 - D_0) - (b + d)N_t]N_t \quad (6.1)$$

The population will be at equilibrium, N^* , when the term in brackets equals 1 because then $N_{t+1} = N_t$. Setting the term in brackets equal to 1 and solving for N to get N^* , which is termed the carrying capacity K , we get, again as for the continuous logistic equation,

$$N^* = K = \frac{B_0 - D_0}{b + d}$$

And defining r as the initial rate of population growth, or $R = B_0 - D_0$, we see that, again as before:

$$K = \frac{R}{b + d} \quad (6.2)$$

The change in population size is the difference between N at $t + 1$ and at t , so subtracting N_t from both sides of the logistic equation gives:

$$\Delta N_t = [(B_0 - D_0) - (b + d)N_t]N_t$$

By substitution of R and K for the terms in $B_0 - D_0$ and $b + d$, we get:

$$\Delta N_t = RN_t \left(1 - \frac{N_t}{K} \right) \quad (6.3)$$

and

$$N_{t+1} = N_t \left[1 + R \left(\frac{K - N_t}{K} \right) \right] \quad (6.4)$$

So far everything looks the same here as before in the derivation of the continuous logistic model, except that the equilibrium is found when the term for the discrete change in N over a time interval (the term in brackets in Eq. 6.1) equals 1 rather than when the derivative, or the rate of continuous change in N , equals 0 as in the continuous logistic model. But, as we shall see, the behaviors of the continuous logistic and discrete quadratic models are not the same even though the two derivations are nearly identical. Before we look at the different behaviors of these two models, let's derive a different version of a discrete logistic model.

The Ricker model

The per capita growth rate at time t of the continuous logistic model can be obtained by multiplying both sides by $1/N$:

$$\frac{dN}{dt} \frac{1}{N} = r \left(1 - \frac{N}{K} \right)$$

Now assume that this per capita growth rate remains constant during a discrete time interval but that it changes in jumps from one time interval to the next. The per capita growth rate during a given time interval is then:

$$r_t = r_0 \left(1 - \frac{N_t}{K} \right)$$

Finally, assume that population growth during a time interval is continuous, and therefore exponential, but with a per capita growth rate that changes discretely. To iterate from N_t one step forward to N_{t+1} , we use the exponential model with r_t as above in place of a constant r :

$$N_{t+1} = N_t e^{r_t} = N_t e^{r_0 \left(1 - \frac{N_t}{K} \right)} \quad (6.5)$$

This is the Ricker model of population growth, used frequently in fisheries biology (Ricker 1954).

The Beverton–Holt model

Finally, we can obtain a discrete time-step version of the continuous logistic model by integrating it forward not from t to T as in Chapter 5, but from t to $t + 1$:

$$N_{t+1} = \frac{K}{1 + \left(\frac{K - N_t}{N_t} \right) e^{-r_0}}$$

Setting $e^{r_0} = \gamma_0$ we obtain after a bit of algebra (Gurney and Nisbet 1998, Turchin 2003):

$$N_{t+1} = \frac{\gamma_0 N_t}{1 + \left(\frac{\gamma_0 - 1}{K} \right) N_t} \quad (6.6)$$

Eq. 6.6 has an identical form to a model proposed by Beverton and Holt (1957), which is also used in fisheries research.

So now we have three different discrete versions of the logistic model. The discrete quadratic model is derived in an identical manner as the continuous logistic model by relaxing the assumption of constant per capita birth and death rates, but beginning with discrete geometric growth rather than continuous exponential growth. The Ricker model is derived by assuming continuous and therefore exponential growth over a time interval, but with the net per capita growth rate changing discretely from one time interval to the next. Finally, the Beverton–Holt model integrates the logistic model for one time step assuming that the initial exponential growth rate for the first time step remains constant.

Do these models behave the same as the continuous logistic model we derived and analyzed in Chapter 5? Do they behave identically to each other? Let's look in detail at the behavior of the discrete quadratic model.

Stability, periodicities, and chaos in the quadratic model

First, as always, we find the equilibrium points where $N_{t+1} = N_t$ for Eq. 6.4. This will happen when $N_t = 0$. It will also happen for values of N_t in which the term in brackets in Eq. 6.4 = 1:

$$1 = 1 + R \left(\frac{K - N_t}{K} \right)$$

which of course happens when $N_t = K$. As for the continuous logistic model, there are therefore two equilibria, 0 and K . So far, we have identical equilibria in both the continuous logistic and the discrete quadratic models. No surprises here.

What about the stabilities of these two equilibria? To examine the stabilities, we again find the eigenvalues and evaluate them at the two equilibria. To find the eigenvalue, we take the derivative of Eq. 6.4 with respect to N_t :

$$\frac{\partial f}{\partial N_t} = \lambda = 1 + R \left(\frac{K - 2N_t}{K} \right) = 1 + R - 2R \left(\frac{N_t}{K} \right)$$

At $N^* = 0$,

$$\left. \frac{\partial f}{\partial N} \right|_{N^*=0} = \lambda_0 = 1 + R - 2R \left(\frac{0}{K} \right) = 1 + R \quad (6.7a)$$

For a discrete model, a perturbation near an equilibrium will grow geometrically (and the equilibrium will therefore be unstable) when $|\lambda| > 1$. Since $R > 0$, $|\lambda|$ near $N^* = 0$ will always be greater than 1 and the size of the perturbation will therefore grow geometrically. Consequently, $N^* = 0$ is locally unstable when $R > 0$ for the discrete quadratic model, just as it was for the continuous logistic model. Still no surprises.

What about $N^* = K$? If K is stable for $N > 1$, then the continuous and discrete models are identical and we are home free. Let's evaluate the eigenvalue at $N^* = K$:

$$\left. \frac{\partial f}{\partial N} \right|_{N^*=K} = \lambda_K = 1 + R - 2R \left(\frac{K}{K} \right) = 1 - R \quad (6.7b)$$

Now comes the surprise: $N^* = K$ is not always stable. The size of the perturbation will decay (and $N^* = K$ will be stable) when $|\lambda| = |1 - R| < 1$. This will happen only when $0 < R < 2$. When $R > 2$, $|\lambda| > 1$ and the perturbation will grow; K is then an unstable equilibrium. The stability of K in the quadratic model therefore depends on the magnitude of $R > 0$, unlike for the continuous logistic model where K was stable for all $N > 0$ when $r > 0$. The discrete quadratic model is not equivalent in all respects to the continuous logistic model.

Because the two models are not equivalent in all respects (equivalent in most doesn't count!), which model is chosen as a mathematical analog of a population needs to be based on the natural history of the organism. For example, organisms with univoltine life cycles require the discrete time version while organisms with overlapping generations, in which reproduction is always occurring in some fraction of the population, require the continuous time version. The Moral of the Story Is: A sound knowledge of natural history is very important in choosing and interpreting the mathematical analog of biological systems. Elegant mathematics does not overcome biological ignorance but only compounds it. On the other hand, elegant mathematics can enhance biological intuition.

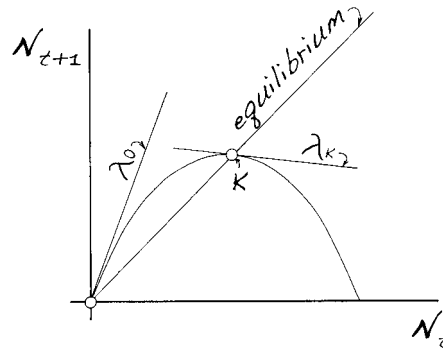
So what happens as R increases, especially when $R > 2$? The linear eigenvalue analysis at the equilibrium points only tells us that the fixed point equilibrium is no longer stable, but it does not tell us what happens next. When linear equations are unstable, they diverge to infinity from all points. When nonlinear equations are unstable, they could also diverge to infinity, but they could also do something else. The "something else" depends on the higher order N^2 in the original quadratic model.

To set the stage for a systematic examination of how the behavior of the discrete quadratic model depends on R , let's recall that in the graph of dN/dt vs. N for the continuous logistic map, as N increases from 0 to K , dN/dt moves smoothly and continuously along this parabola until it reaches K . All the (real-valued!) points along the parabola are covered in sequence.

This is not the case for the discrete quadratic model. First, the solutions for any population size change in discrete steps. Second, they do not increase to K in sequence when $R > 2$. Let's see why.

To proceed, we recall that an equilibrium is a value of N for which $N_{t+1} = N_t$, which is simply a straight line with a slope of 1. We then plot both Eq. 6.4 and the

Fig. 6-1 Location of equilibria in the discrete quadratic model.



equilibrium line $N_{t+1} = N_t$ on the same graph (Fig. 6-1). The equilibrium points of the discrete quadratic model, 0 and K , happen at its intersections with the equilibrium line. As always, the eigenvalues are the slopes of the lines tangent to the function at 0 and K .

Algebraically, we would iterate the quadratic model forward by beginning with N_t , calculating N_{t+1} , and then use N_{t+1} as the next value of N_t . Graphically, this is equivalent to beginning at some point on the N_t axis, moving vertically until we reach the graph of the discrete quadratic model to get a value for N_{t+1} , then moving horizontally to the line $N_t = N_{t+1}$. This procedure graphically converts N_{t+1} to the next N_t . We then move from this point vertically up or down to the quadratic model to get the next N_{t+1} and repeat.

Let's see what happens when $-1 < R < 0$. In this case, $0 < \lambda_0 < 1$ and we get something like geometric decay of population density towards a stable $N^* = 0$ (Fig. 6-2).

But when $0 < R < 1$, then $\lambda_0 > 1$ and $N^* = 0$ is unstable, so the population grows away from 0 towards K in discrete jumps (Fig. 6-3). But $0 < \lambda_K < 1$ so the difference between N_t and K decreases in jumps of ever decreasing size. So as R passes through 0, the two equilibrium points switch stabilities resulting in a transcritical bifurcation, exactly the same bifurcation in the continuous logistic model at the same critical value of R . When $-1 < R < 0$, the trajectory converges on zero, but when $0 < R < 1$, the trajectory converges steadily on K .

Let's increase R further. When $1 < R < 2$, then $-1 < \lambda_K < 0$, and N_t approaches K with damped oscillations (Fig. 6-4). By the way, these plots are called cobweb plots (the reason for the term is especially evident in this drawing) and cobwebbing is the doodling that creates them (Strogatz 1994). This is an excellent doodle to while away

Fig. 6-2 Cobweb plot of the discrete quadratic when $-1 < R < 0$ and $\lambda_0 < 1$.

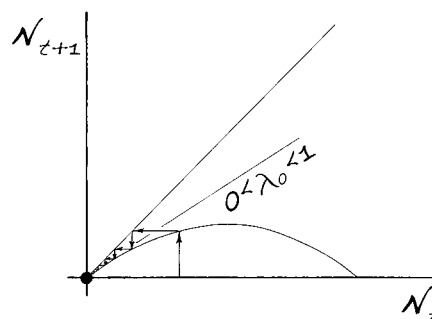


Fig. 6-3 Cobweb plot of the discrete quadratic when $0 < R < 1$ and $\lambda_0 > 1$.

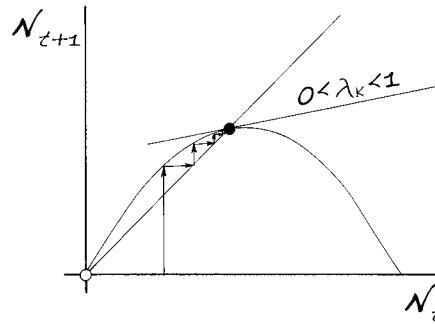
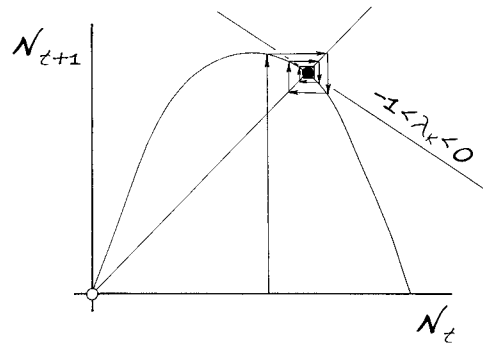


Fig. 6-4 Cobweb plot of the discrete quadratic when $1 < R < 2$ and $-1 < \lambda_K < 0$.



those idle moments in airports, train stations, faculty meetings . . . Let's have some fun and practice cobwebbing before we go further.

Exercise 6.1

Sketch any old function on graph paper and have it intersect the equilibrium line at more than one point (it doesn't matter if you don't know the equation – just sketch a function). Cobweb the function beginning with points near each of the equilibrium points. How does the cobweb plot depend on the direction(s) of the slope(s) of the function at its intersection(s) with the equilibrium line? Now change the slope of the function at each equilibrium point and cobweb the new function. What happens when the slope gets steeper? Do this before you read further – no fair peeking at the answer!

What happens when $R > 2$? As we have seen, the eigenvalue at K is then less than -1 and the solution no longer converges on K , which is now unstable. Instead, the solution oscillates between two points, one greater than and the other less than K , with a period of two time steps (Fig. 6-5). If we plot N_t against t , we can see the population fluctuations more clearly (Fig. 6-6). The period of oscillation is two time steps because starting from either point, it takes two steps to get back to it: the first time step puts you on the other point and the next one gets you back. You can think of the model as orbiting from one point to the other. This is called a period-2 cycle or an orbit of period 2. In a sense, a stable K is a period-1 orbit because the orbit stays on K after each time step.

These two points are equilibria in period 2 because, once the solution is on either of them, it returns to it in two time steps. A point is an equilibrium of a discrete model in period- τ if the value at t is equal to the value at $t + \tau$, τ being the lag in

Fig. 6-5 Cobweb plot of the discrete quadratic when $2 < R < 2.45$ and $\lambda_K < -1$.

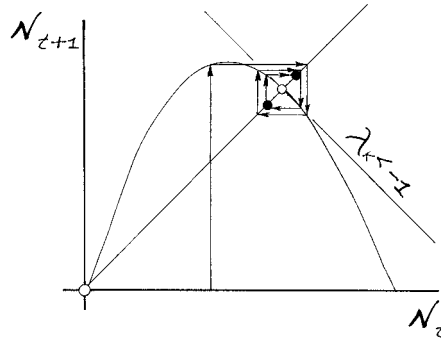
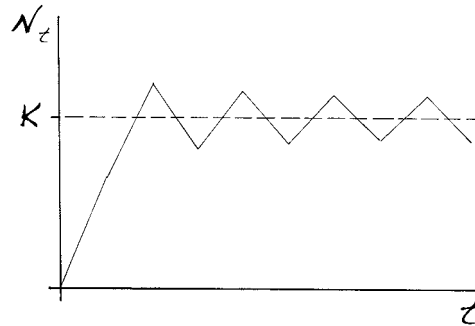


Fig. 6-6 Population oscillations in the discrete quadratic when $2 < R < 2.45$.

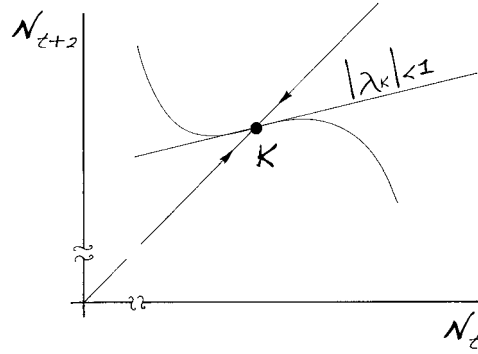


time step units. Pretend you are inventorying a population that obeys the quadratic model with a value of R slightly greater than 2 and which is currently oscillating between these two points. You inventory the population density N_t with a lag every two time steps ($\tau = 2$; lack of funds prevents you from doing inventories more frequently). What do you see? You see the population density always equal to one of these two equilibrium values. (If you inventoried the population on the off-time steps, you would always see the population density at the other equilibrium.)

Therefore, $R = 2$ is a critical value producing a qualitative change in the behavior of the model near the equilibrium point K , just like $r = 0$ was also a critical value producing a transcritical bifurcation in the continuous logistic model. Only now this new critical value of R does not result in two solutions switching stability, as in the transcritical bifurcation, but the appearance of two new equilibria that the model flips between combined with the old equilibrium in between them becoming unstable. When a stable equilibrium like K becomes unstable at a critical value of a parameter or combination of parameters and spawns two new equilibria that the model oscillates between, we have a new kind of bifurcation, called a flip or period doubling bifurcation (Strogatz 1994), which is closely related to the Hopf bifurcation that gives rise to stable limit cycles in predator–prey differential equation models (May 1978), which we shall examine later.

Moreover, these two new equilibria are locally stable when $\tau = 2$. We can see why $R = 2$ causes K to become unstable: because $|\lambda| > 1$. But *why* does an instability of K at $R = 2$ result in two other locally stable equilibria to either side of it when $\tau = 2$? To see this, let's derive an equation that gives N_{t+2} from N_t and examine the

Fig. 6-7 Behavior of the discrete quadratic two time steps into the future when $R < 2$.



equilibrium points of that equation and their stability. First, we write the equation for N_{t+2} in terms of N_{t+1} :

$$N_{t+2} = N_{t+1} + RN_{t+1} \left(1 - \frac{N_{t+1}}{K} \right)$$

Then substitute Eq. 6.4 for N_{t+1} , resulting in this rather formidable-looking beast:

$$N_{t+2} = \left[N_t + RN_t \left(1 - \frac{N_t}{K} \right) \right] + R \left[N_t + RN_t \left(1 - \frac{N_t}{K} \right) \right] - \frac{R}{K} \left[N_t + RN_t \left(1 - \frac{N_t}{K} \right) \right]^2 \quad (6.8)$$

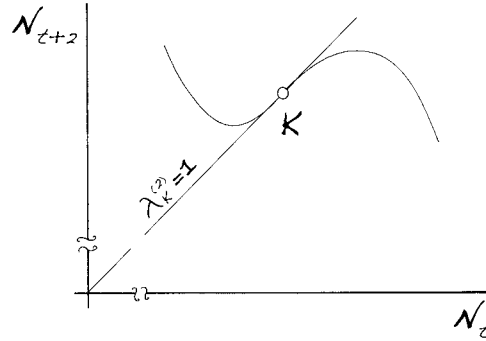
Now this equation is quartic in N_t (if you square the last term in brackets one of the terms will be N_t^4). The first term on the r.h.s. gives the value for N_{t+1} and the last two terms give corrections for projecting N_{t+1} forward to N_{t+2} .

It is these higher order terms, especially the quartic term in N_t , that control what happens after K becomes unstable and then gives rise to the two new equilibria when $R > 2$. Quartic equations contain an S-shaped twist. This particular equation has such a twist centered on K . Now let's plot N_{t+2} vs. N_t . The line $N_{t+2} = N_t$ is the equilibrium line for $\tau = 2$. When $R < 2$, the equation crosses the equilibrium line once, at K (Fig. 6-7). Since it is crossing the equilibrium line at K with a slope (eigenvalue) whose absolute value is < 1 , K is a stable equilibrium not only for a lag of 1 (N_{t+1}) but also for a lag of 2 (N_{t+2}). That is, the trajectory will remain on K not only for the next time step, but for the one after that (and after that, and . . .). You can verify that geometrically by placing arrows on the equilibrium line to either side of K according to the value of the slope relative to the equilibrium line, as we did in Chapter 5 on the x -axis for the continuous logistic equation. The x -axis was, as you recall, the equilibrium line for $N(t)$ on the graph dN/dt vs. $N(t)$.

Can we find an algebraic expression for the eigenvalue at K for Eq. 6.8? This will tell us if K is a stable equilibrium after two time steps. Well you could differentiate this beast with respect to N_t and then solve it for $N_t = K$, but there is an easier way. Recall that at the equilibrium K , N_{t+1} is exactly given as:

$$N_{t+1} = \lambda_K^{(1)} N_t$$

Fig. 6-8 Behavior of the discrete quadratic two time steps into the future when $R = 2$.



The superscript (1) on λ indicates that we are talking about the eigenvalue of Eq. 6.4, the single-time step model ($\tau = 1$). Now,

$$N_{t+2} = \lambda_K^{(1)} N_{t+1} = \lambda_K^{(1)} \lambda_K^{(1)} N_t = (\lambda_K^{(1)})^2 N_t = \lambda_K^{(2)} N_t$$

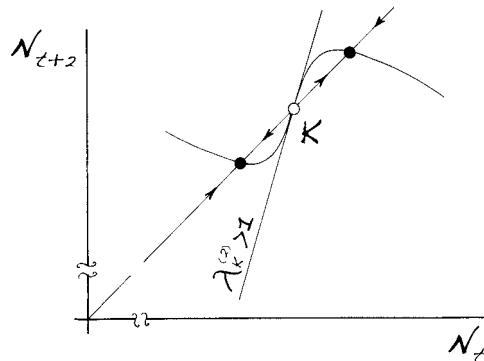
So the eigenvalue at K of the two-time step model, which we have indicated with $\lambda_K^{(2)}$, is the square of the eigenvalue, $\lambda_K^{(1)}$, of the single-time step model (May 1978). What is $\lambda_K^{(2)}$? It is:

$$\lambda_K^{(2)} = (\lambda_K^{(1)})^2 = (1 - R)^2 = 1 - 2R + R^2$$

Now when $R = 2$, $\lambda_K^{(2)} = 1$. At K , the slope of Eq. 6.8 then is equal to the equilibrium line. The two equations are therefore tangent to each other at K when $R = 2$ (Fig. 6-8). Increasing R has twisted Eq. 6.8 a bit around the pivot point K .

What happens when we increase R further? Then the twist increases still more, causing Eq. 6.8 to cross the equilibrium line at three points (Fig. 6-9). The two new equilibrium points have now popped into existence at equal distances to each side of K . Now draw arrows on the equilibrium line indicating the direction of change of the population density to either side of K and the two new equilibria. You can see that after a small perturbation to either side of K , the population moves towards two new equilibria, landing on alternate sides of K for each single time step as it migrates

Fig. 6-9 Behavior of the discrete quadratic two time steps into the future when $2 < R < 2.45$.



out to them. Either of the two new equilibria is stable *after every two time steps* (in the intervening time step, the population has jumped to the other equilibrium, but then jumps back in the next step). So long as the absolute value of the slope of Eq. 6.8 at the two new equilibria is less than 1, these equilibria will be stable in two time steps. (Demonstrate this for yourself by cobwebbing the function near the two new equilibria in this drawing. Enlarge it with a copier if you need to.) The trick to understanding all this is to keep in mind whether you are talking about stability of a point after a single time step (K when $R < 2$) or after a period of two time steps (the two new equilibria when $R > 2$). A period-2 cycle has two equilibria which are stable after two time steps ($\tau = 2$).

If R increases further from 2, the two solutions begin to move away from K and the amplitude of the period increases. But this does not continue indefinitely. If we graphed the function for N_{t+4} we would see that there is another set of S-twists for a lag equal to four time steps and which are centered on the two equilibria to either side of K . As R increases, the slopes of those twists for $\tau = 4$ also steepen. Eventually, at $R = 2.45$ two new equilibria framing each of the two previous equilibria (which in turn framed K) pop into existence, exactly as what happened for K when $R > 2$ in Eq. 6.8. Now the two, period-2 equilibria to either side of K have become unstable and the four, period-4 equilibria to either side of them are stable. Both the number of solutions and the number of periods of the cycle have doubled. This phenomenon is called period doubling of the cycle. The rule is that each of the four points must be visited once before any of them can be revisited to close the orbit.

This doubling of solutions and periods at critical points now begins to happen very quickly as R increases further. New S-twists form, popping matched pairs of equilibria into existence once some critical value of R is exceeded and generating more orbits of periods which are double that of the previous orbit. A period-8 cycle is spawned at $R = 2.544 \dots$, a period-16 cycle at $R = 2.564 \dots$, a period-32 cycle at $R = 2.5687 \dots$, and so forth. All previous orbits of shorter periods become unstable. An intricate set of embedded twists forms, each centered on an equilibrium point, producing more and more oscillations at longer and longer periods. In general, period doubling produces a period τ cycle which has τ equilibria which are stable after τ time steps as the value of a control parameter is increased and as critical values are exceeded. But the mechanism, the formation of S-twists centered in previous equilibria, looks the same at all lengths of the lags. This is the famous fractal structure of this model (Strogatz 1994, Hirsch et al. 2004), fractal because the structure (S-twists) looks and behaves the same at all time scales.

When $R = 2.57$, the number of periods and equilibrium points is infinite. The time series of N_t now looks random, *but there are no random numbers in Eq. 6.4!* Here is what it looks like for $R = 3$, $K = 1$, and $N_0 = 0.1$ (Fig. 6-10; the first 50 points have been discarded so that initial transients settle down). The discrete quadratic model with $R > 2.57$ is not random but completely deterministic because every successive point, N_{t+1} , is determined precisely and exactly by the values of N_t , R , and K with no exogenous random number. Underlying this erratic sequence of values of N_t , there is still the exact parabolic structure of the deterministic quadratic model. Every value of N_t is somewhere on the parabola of the quadratic model, only now successive points (N_{t+1}) bounce all over the place rather than appear nicely in sequence as in the continuous logistic model. Here are the first 10 points of the sequence in the above drawing plotted in their positions on the parabola of N_{t+1} vs. N_t (Fig. 6-11).

Fig. 6-10 Chaotic behavior of the discrete quadratic when $R > 2.57$.

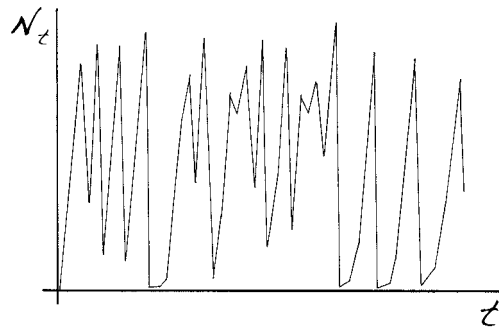
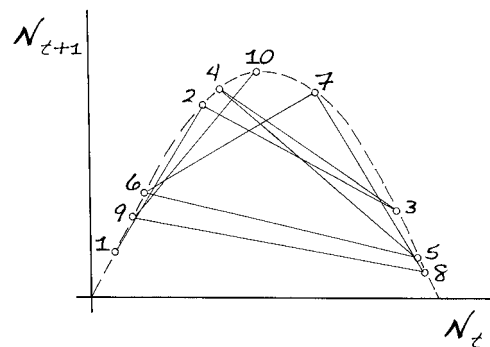


Fig. 6-11 Successive points on the parabola of the discrete quadratic are visited without any pattern when $R > 2.56$.



This erratic time sequence of population densities produced by an underlying deterministic structure at some critical values of one or more parameters is called *deterministic chaos*, or *chaos* for short. The parabola is an invariant because the trajectories all end at some point or points on it in some fashion: either stable fixed-point equilibria (remember K is an attractor when $R < 2$) or in sets of points that define periodic orbits, which are also attractors. But when the trajectory is chaotic, every one of the infinite number of points on the parabola is visited without any of them ever being revisited: there are an infinitely dense set of unstable periodic orbits which never get to repeat exactly as $t \rightarrow \infty$ (Fig. 6-11). This is a strange sort of attractor, indeed. In fact, it is called a *strange attractor*. A strange attractor is sometimes called aperiodic because no value is ever repeated during a trajectory. Instead, the trajectory comes infinitesimally close to repeating any single value. This is not because the trajectory is bouncing randomly between points, but because the trajectory gets passed from one unstable periodic orbit to another for an infinite number of periodic orbits. No orbit can be repeated before all orbits are visited and because the number of orbits is infinite, none of them gets repeated as $t \rightarrow \infty$.

Note also that for the above sequence of points, some come perilously close to 0, or extinction of the population. In fact, for $R > 3.0$, extinction in the discrete quadratic model is inevitable. There appears to be an upper limit on population growth rate in this model, a point to which we shall return shortly.

Another necessary feature of deterministic chaos is its sensitive dependence on initial conditions. That is, two solutions which are initially very close will not remain very close but will eventually separate or diverge from one another at an exponential rate. Strogatz (1994) uses the metaphor of folding of pastry dough or kneading bread

to demonstrate this: As you stretch and fold a mass of dough, turn it 90 degrees, stretch and fold again, two points that may have started out next to each other very quickly become stretched far apart.

Try it next time you bake bread – stick two raisins in the dough about 2–3 cm apart, and knead the dough by stretching, folding, and rotating 90 degrees. Repeat this process about 150 times to get a long enough time series and also a nice, dense bread. How far apart are the raisins now? If you can't see them, cut the dough in slices whose thickness was half the original distance between the raisins. Are the raisins in the same slice? If not, then they have separated during the stretching, folding, and rotating of the kneading process. If they are in the same slice, how far apart are they? This stretching and folding of real dough is geometrically analogous to the S -twisting of Eq. 6.4 in longer lags of time steps as R increases, which caused period doubling bifurcations.

The higher order terms in Eq. 6.8 that produced the S -twists and period doublings also causes points that initially were arbitrarily close to separate. Restate Eq. 6.8 for another initial condition $N = N_0 + \varepsilon$, arbitrarily close to N_0 , for just the first two time steps:

$$\begin{aligned} N_{t=2} = & \left[(N_0 + \varepsilon) + R(N_0 + \varepsilon) \left(1 - \frac{N_0 + \varepsilon}{K} \right) \right] \\ & + R \left[(N_0 + \varepsilon) + R(N_0 + \varepsilon) \left(1 - \frac{N_0 + \varepsilon}{K} \right) \right] \\ & - \frac{R}{K} \left[(N_0 + \varepsilon) + R(N_0 + \varepsilon) \left(1 - \frac{N_0 + \varepsilon}{K} \right) \right]^2 \end{aligned}$$

You can see that even after the first two time steps, the value of the initial separation, ε , will have it's own dynamics because there will be terms such as $R\varepsilon^2$, $R^2\varepsilon^2$, $R^3\varepsilon^4$, $R^3\varepsilon N_0$, etc., each with it's own sign. When R is large (>2.57), these higher order terms eventually cause ε to grow. As ε grows, the solutions separate from their initial difference.

Sensitivity to initial conditions can be mathematically described more generally in this way. Start with some initial population density, N_0 , then add a small amount, ε , to this, producing another initial condition $N_0 + \varepsilon$, arbitrarily close to N_0 . You can think of ε as a tiny perturbation at t_0 , producing another trajectory. As the trajectories are produced by iterating the model from the two initial conditions many times (i.e., as $t \rightarrow \infty$) ε will either decay or grow exponentially:

$$|\varepsilon_n| = |\varepsilon_0| e^{\Lambda n} \quad (6.9)$$

where ε_0 is the initial separation between N_0 and $N_0 + \varepsilon$ and ε_n is the separation after n iterations. The coefficient of the exponent, Λ , is called the Lyapunov exponent and measures the rate of change of the separation of the two trajectories. If Λ is greater than 0, then the separation grows and the two initial values become far apart. Therefore, the predictions of the model after a period of time are exponentially sensitive to differences in initial conditions. Eventually the continued growth of the separation

ends when the trajectories converge on the strange attractor, so the instability due to chaos does not mean unbounded behavior: it is bounded by the strange attractor. Positive Lyapunov exponents are characteristic of chaotic systems. It is important to recognize that the initial difference can be arbitrarily small and still the points will eventually separate. Therefore, separation of the two trajectories cannot be offset by more precise measurements of initial conditions, although the separation may be initially very small for the first several time steps or more. Better measuring devices do not always help!

More formally (Strogatz 1994), we can divide both sides of Eq. 6.9 by $|\epsilon_1|$ and taking the $\ln$ of both sides, solve for Λ :

$$\Lambda = \frac{1}{n} \ln \left| \frac{\epsilon_n}{\epsilon_0} \right| = \lim_{\epsilon_0 \rightarrow 0} \frac{1}{n} \ln \left| \frac{f^n(N_0 + \epsilon_0) - f^n(N_0)}{\epsilon_0} \right| \cong \frac{1}{n} \ln |(f^n)'(N_0)| \quad (6.10)$$

where f^n is the n^{th} iteration of a function and $(f^n)'(N_0)$ is the derivative of the n^{th} iteration with respect to N evaluated at N_0 . This derivative is, of course, exactly the eigenvalue for successive iterations which we used to determine the stability of the new equilibria which pop into existence with period doubling bifurcations. There is, therefore, a deep connection between Lyapunov exponents and eigenvalues (see also, for example, Strogatz 1994 and Hirsch et al. 2004).

One way to get a rough and ready statistical estimate of Λ is to generate two time series of solutions of a function starting from two different, but nearby, initial conditions. After the time series are generated for a sufficiently long time, subtract the value of one function from the other at each iteration and take the absolute value. This gives you a series of the separations, ϵ_n . Then, using the $\ln$ transform of Eq. 6.9,

$$\ln |\epsilon_n| = \ln |\epsilon_0| + \Lambda n$$

regress $\ln |\epsilon_n|$ against n . The slope of this regression is an estimate of Λ . It is probably best to do this several times for several different pairs of time series and find the expected value of Λ (the mean if the population of estimates of Λ is normally distributed). These estimates of Λ are, however, very sensitive to departures of the residuals of the regression from normality and the other assumptions of regression analysis, so if your estimate of Λ is close to the critical value of 0, you should treat it with caution. More sophisticated methods for estimating Λ from time series of model output or especially from data which may be contaminated by random noise are also available (Ellner and Turchin 1995).

Chaos in dynamical systems was first noticed by Henri Poincaré in an analysis of stability in a three-body model of the solar system, which is a great story involving the sixtieth birthday celebration of the King of Sweden, a gold medal, and a race to fix a mistake in a proof (Diacu and Holmes 1996, Stewart 2002, Mawhin 2005). Further developments in mathematics and theoretical meteorology (recounted in Lorenz 1996) followed. But it was the papers of May (1974, 1975a, 1976) and May and Oster (1976) which brought the attention of ecologists to the possibility of simple underlying causes for complicated behaviors such as periodicities and deterministic chaos.

In nature, what we often see is a complicated time series of changes in population densities, including oscillations at regular periods. We all have the temptation to explain such complicated dynamics by complicated or periodic driving variables. These may indeed be the causes, but as the title of May (1976) and the above example show, complicated dynamics need not arise from complicated or even external causes.

So do natural populations behave chaotically? The difficulties of separating deterministic chaos from random noise in real data are formidable (Ellner and Turchin 1995), so this is not an easy question to answer. Constantio et al. (1997) have induced chaotic dynamics in a flour beetle population and Becks et al. (2005) have induced chaotic dynamics in a predator–prey system in a chemostat. Both of these, as well as other recent experiments, are performed under highly controlled laboratory conditions which reduce contaminating random noise.

In wild nature, both random noise and deterministic chaos may be operating and the problem of separating noise from possible deterministic chaos is formidable (Ellner and Turchin 1995). Arguments for (e.g., Schaffer 1984, Schaffer and Kot 1985, 1986) and against (e.g., Berryman and Millstein 1989) chaos, therefore, are based largely, although not entirely, on theoretical arguments. The arguments for chaos are the obvious existence of complicated population cycles, some of which could be pushed to chaos by increases in population growth rate. The arguments against chaos are that values of R near 3.0 are not very common and, as we have seen, populations with $R > 3.0$ go extinct. Therefore, any population that behaves chaotically has a high probability of extinction during one of its very large oscillations, so if chaos happens it is quickly selected against.

Nonetheless, there is some evidence of chaotic populations in nature (Shaffer 1984, Turchin 1993, Ellner and Turchin 1995). Ellner and Turchin (1995) analyzed long time series of population inventories for 28 datasets and found Lyapunov exponents slightly greater than and statistically different from 0 in a few populations, mainly Lapland voles and possibly a few other northern populations. This seems to agree with the findings of Turchin (1993) that chaotic population dynamics in small mammals are more likely at higher latitudes and the analysis by Shaffer (1984) of a strange attractor in lynx fur return record of the Hudson Bay Company. These species all have high population growth rates with R often considerably greater than 2 under favorable conditions of abundant food supply and low predation. In addition, the strongly seasonal boreal and polar environment essentially creates a discrete time series of growth and reproductive seasons that are widely separated with an intervening period of high mortality during winter, thus closely mimicking a discrete time-step environment. Turchin (2003) gives additional examples of recent investigations of chaos in wild populations.

Regardless of whether populations in the wild are chaotic in the strict sense, the value of investigating the complex behavior of simple models that yield population oscillations of multiple periods through period doubling bifurcations when a parameter crosses a critical value have stimulated research into the possibility of other bifurcations in ecological models and the possibility that complicated behavior could be explained by simple underlying causes, such as crossing critical values of one or more parameters. We shall meet more bifurcations in the next chapter and in later chapters as well.

Exercise 6.2

Analyze the Ricker and/or the Beverton–Holt models either graphically using cobwebbing or analytically by calculating eigenvalues. Do these also have critical points where period-doubling occurs? A spreadsheet, which is simply a difference equation solver, may be useful to graph the models for different values of R .

Feedbacks and time lags in ecological systems: the biological mechanisms producing oscillations and chaos

We have uncovered the mathematical reasons for the appearance of oscillations and chaos in the discrete quadratic model, but we still have not addressed the biological reasons. What biological mechanisms are suggested by our mathematical analyses of Eq. 6.4 that could cause oscillations and chaos in natural populations? There are two: internal feedbacks in population regulation and time delays in the population response to the feedback.

The feedback is implicit in the density-dependent regulation of per capita birth and death rates through which we derived the logistic and discrete quadratic models. The time delay is implicit in the discrete nature of difference equation models. N does not change instantaneously, as for continuous differential equation models, but only after a discrete time step of finite length. That is, the feedback from N_t to N_{t+1} through the underlying changes in birth and death rates happens only at the end of one time step and the beginning of another. It is well known mathematically that delays can cause solutions of nonlinear equations which contain implicit feedbacks to oscillate (Jones 1962). The combination of a negative feedback (density-dependence) with a time delay in the discrete quadratic model causes the appearance of oscillations, especially as R (and therefore the strength of the feedback) increases.

To see intuitively why delays in a natural feedback system can cause oscillations, consider a feedback system which most of you (at least in cold climates) are familiar with: the thermostat–furnace system that (you hope!) maintains a constant temperature in your home. When the air cools, a bimetal strip in the thermostat bends and eventually makes contact with an electrical junction, causing a current to flow to the furnace and ignite the flame. As the furnace heats the air in your home, the bimetal strip bends back, the electrical contact is broken, and the furnace turns off. But if the bimetal strip is too stiff, it makes contact with the electrical junction very slowly and only after the air temperature has fallen by more than is desirable; it also breaks contact only after the furnace has warmed the air more than needed. The stiffness of the bimetal strip introduces a delay in its response that causes oscillations in air temperature – the stiffer the strip, the longer the delay in its response, and the greater the oscillations in air temperature.

Can we make the continuous differential equation of logistic growth oscillate by introducing a delay? The answer is yes, as was first shown by Hutchinson (1948) and analyzed further by Wangersky and Cunningham (1957):

$$\frac{dN}{dt} = rN \left(1 - \frac{N(t - \tau)}{K} \right) \quad (6.11)$$

This delay can be caused, for example, by a starving animal surviving without food for a length of time τ . You can see more clearly how $N(t - \tau)$ delays the growth of $N(t)$ by writing Eq. 6.11 in this way:

$$\frac{dN}{dt} = rN - \left(\frac{rN(t - \tau)}{K} \right) N$$

The rate of population change, dN/dt , is a result of a tension between rN pushing the population onward and $-rN(t - \tau)/K$ pulling it backward.

Recall that $1/|\lambda|$ is the characteristic time scale of change in the perturbation for return to some fraction $1/e$ of its original departure from the stable equilibria (Chapter 2, Eq. 2.16). If the delay, τ , is much longer than $1/|\lambda|$, then the equation will begin to oscillate around a formerly stable equilibrium (May 1978). For the logistic equation, $|\lambda| = r$. Therefore, if $r\tau < 1/e$, the solution will return exponentially to K ; if $1/e < r\tau < 1/2\pi$, it will return with damped oscillations; but if $r\tau > 1/2\pi$, then a stable limit cycle centered on K is spawned (May 1978). The period of this oscillation is approximately 4τ , which May (1978) explains intuitively as follows: a population with density near K and whose delay is long enough to cause a periodic solution will continue to grow past K until the lagged population density $N(t - \tau)$ “catches up” and equals K . Then, dN/dt in Eq. 6.10 will equal 0. This period of “catch up” necessarily takes a length of time τ . The population growth will then become negative as $N(t - \tau)$ overshoots K . Reaching K on the downturn takes another τ time steps, whereupon K is overshoot on the underside, taking 2τ more time steps to return to K at the same slope as we began (return to both the same density and same sign and magnitude of the derivative are required for a recovery to the same initial state). For a population growth rate $r = 2.0 \text{ yr}^{-1}$, a delay of $\tau = \pi/2r = 0.7854$ will initiate oscillations with periods of 3 to 4 years. Such population growth rates are typical for small mammals in northern climates, such as lemmings and other microtines (Pitelka 1967). Moreover, this time delay is on the order of 9 months, remarkably close to the length of winter between successive summers in boreal and tundra regions, during which the lemmings and microtines reproduce (May 1978). Eq. 6.11 is the simplest example of oscillatory dynamics induced by delays in continuous differential equation models with feedbacks.

Summary: what have we learned?

Although there were some differences in the forms and behavior of continuous linear differential equation models and the analogous discrete time-step linear model, the continuous exponential model was the limit of the discrete geometric model as the time step of the latter was reduced to zero. Therefore, with some care to details such as the relation between the exponential r and the geometric R , the same conclusions are reached from either continuous or discrete versions of linear models. Not so for nonlinear models: even when they are derived through identical procedures, the discrete quadratic model has qualitatively different behaviors than the analogous continuous logistic model. These new behaviors began with K first becoming an unstable equilibrium as R increased beyond a critical point and spawning at first two new equilibria, then continued with a rapid series of period doubling bifurcations leading to increased density of equilibria and eventually deterministic chaos. Complicated dynamics, including oscillations as well as chaos, need not be caused by complicated external forcings, but could arise from strong internal forcings (large R) and delayed feedbacks via density-dependence because of the implicit delay in a discrete time-step model. Introducing explicit delays in differential equation models also causes their solutions to oscillate. Therefore, care must be taken in matching the life history characteristics of an organism (univoltine vs. continuous reproduction, age-related delays in producing offspring, etc.) to the particular mathematical form of a model.

Open questions and loose ends

Complex dynamics in discrete versions of generalized logistic models – In Exercise 5.1, you derived another logistic-like model by making per capita birth and death rates depend linearly on the number of pairs of individuals, thereby modeling pair-wise interference in birth and death rates. This resulted in a cubic differential equation model. Discrete versions of this model can also be derived by the methods introduced here for the discrete quadratic model. Do oscillations and chaos appear in this cubic model as R increases? By what mechanism? It would be well worth exploring how different ways of introducing density-dependent birth and death rates results in periodicities, bifurcations, and chaos in the family of generalized logistic models. Do the complex behaviors depend in some predictable and understandable way on how density-dependence is introduced?

Time delays in population models – Explicit time delays can produce some wild dynamics in both continuous and discrete models (see Kot 2001, pp. 70–91 for an introduction and references and MacDonald 1978, 1989 for a fuller treatment), including homoclinic and other sorts of bifurcations that we have not the space to consider here. Many of these models have the time delay “put in by hand” as in the insertion of $N(t - \tau)$ for $N(t)$ in Eq. 6.10, rather than have the time delay emerge from an explicit mathematical description of a specific biological mechanism. For example, Hutchinson (1948), who introduced Eq. 6.10, suggested that the delay might be accomplished by delayed seed germination, but this mechanism is not explicit in the equations, merely an interpretation of what Eq. 6.10 might depict. To further our understanding of how time delays arise and operate in the real world, we need to derive models from explicit assumptions of how time delays affect per capita birth and death rates or other specific and explicit biological processes that affect population growth, not simply put in the delay by hand. Time-delay models derived explicitly from processes hypothesized to be responsible for the delay can be tested against experiments in which the delay mechanisms can be manipulated to see if periodicities of the correct amplitude and frequency really do arise. This would be a more satisfying test of the role of time delays than the coincidence of period-4 cycles in the delayed logistic model with a time delay of 9 months with similar behaviors in real populations of northern microtine rodents.

Control of chaos – If R represents the net growth rate of a population and if the underlying biological processes of mating, birthing, and surviving have at least partial genetic control, then there would appear to be a selection for genes or genetic systems that increase reproductive potential and increase R . But, as we have seen, increasing R indefinitely can also destabilize trajectories of population densities, eventually causing extinction with high values of R . Is there some way to control the onset of chaos, either through natural processes or through intentional intervention via management?

Control of chaos means reversal of period-doubling bifurcations as a parameter such as R increases (Stone 1993, Solé et al. 1999). Period doubling reversal can be achieved either by small perturbations to a control parameter such as R at the period which is to be prevented (Ott et al. 1990) or by small perturbations to the state variable such as population density, also applied at a selected period (Güemez and Matías 1993, Parthasarathy and Sinha 1995). For example, control of period-doubling in the discrete quadratic model could be accomplished through the Ott et al. (1990) method by pulsing R at some period (Solé et al. 1999):

$$N_{t+1} = N_t + (R + \gamma)N_{t-(p-1)} \left(1 - \frac{N_{t-(p-1)}}{K}\right)$$

where R is decreased by some amount γ every p generations, p also being the periodicity which is to be removed. Since $R = B_0 - D_0$, this control could be accomplished by either decreasing B_0 or increasing D_0 at every p^{th} generation through natural selection (Doebeli 1993) or through pulsed forcings by periodicities in external environmental factors such as climate (Doebeli and Ruxton 1997).

Alternatively, period doubling can be controlled by augmenting or reducing population density with period pulses every p^{th} generation (Solé et al. 1999):

$$N_{t+1} = N_t + RN_{t-(p-1)} \left(1 - \frac{N_{t-(p-1)}}{K}\right) (1 + \gamma)$$

where γ is a proportional increase or decrease in population density applied every p^{th} generation (Güemez and Matías 1993). Alternatively, we can introduce γ as a pulsed increase or decrease in population density applied every p^{th} generation:

$$N_{t+1} = N_t + RN_{t-(p-1)} \left(1 - \frac{N_{t-(p-1)}}{K}\right) + \gamma$$

(Parthasarathy and Sinha 1995). The control parameter γ could represent pulsed stocking or harvesting in a managed population or pulsed immigration or emigration in a natural population (McCallum 1992, Stone 1993, Solé et al. 1999).

Of course, γ can take on any reasonable functional form which represents different harvesting or stocking strategies or different immigration/emigration dispersal dynamics. How does reversal of period doubling depend on the functional form of γ ? Are other types of bifurcations introduced with different functional forms of γ ? We will investigate the effect of different forms of per capita harvesting functions on behavior of the continuous logistic model in the next chapter, but their applications to control of chaos in the discrete quadratic model or other discrete models of population growth are problems rich in possibilities.

7 Harvesting and the logistic model

Now that we have a reasonable model of a single population, namely the logistic model, we can begin to ask how this population might interact with another population. We will assume that the second population harvests or consumes a portion of a resource or prey population which otherwise grows according to the logistic model in the absence of harvest. For the moment, we will ignore the dynamics of the harvester/consumer and assume that it can harvest at will according to some per capita strategy and its population density.

A general form for such a harvested population is:

$$\frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{K}\right) - f(N_1)N_2 \quad (7.1)$$

where $f(N_1)$ is the per capita harvest rate of N_1 by individuals of N_2 and N_2 is the population density of the harvester or consumer. In this chapter, we will treat N_2 as a parameter. In succeeding chapters, we will allow N_2 to have its own population dynamics as a function of both its own density as well as that of N_1 , thus coupling the two populations into a single dynamic system.

We must now either derive or pose some reasonable forms of $f(N_1)$ and determine their effect on the equilibria of Eq. 7.1 and their stability. We will especially examine how these equilibria and their stability differ from the unstable equilibrium $N_1^* = 0$ and the stable equilibrium $N_2^* = K$.

Before proceeding, I wish to offer a few observations and comments on terminology. In the ecological literature, the per capita harvest strategy, $f(N_1)$, is commonly called a consumer's "functional response" and the total consumption $f(N_1)N_2$ is commonly called the consumer's "numerical response" to the resource or prey consumed. These terms have always seemed rather vague and confusing to me. After all, anytime we pose or derive a function to describe how Y responds to X , we get a "functional response" and if we wish to use the natural or real number systems to quantify the magnitude of the response, then we have a "numerical response." It seems to me that the terms "per capita harvest strategy" or simply "harvest strategy" for $f(N_1)$ and "consumer (or harvester) population response" for $f(N_1)N_2$ are more precise descriptions of what is usually meant by "functional" and "numerical" responses.

By "harvest" and "harvest strategy" I do not imply any teleological or anthropomorphic intent on the part of the consumer. The term "harvest" can equally well apply to the uptake of a nutrient by a plant (the plant "harvests" the nutrient from the soil), the consumption of a plant by an herbivore or an herbivore by a predator, or the harvest of a resource by an industry or management agency. Indeed, we will

see that one of the more commonly employed “functional responses” can be derived equally well from either considerations of biting and processing a prey or from considerations of the rate of delivery of nutrients to a root and the root’s ability to process them once they arrive. Therefore, at the risk of bucking a well-established tide, for the remainder of this book I will use the term “harvest strategy” to refer to the per capita harvest rate of a resource by a consumer instead of the term “functional response” and the term “consumer population response” instead of the term “numerical response.”

Let’s now derive some forms of harvest strategies and see what their consequences are for single populations undergoing logistic growth in the absence of harvest. Beddington and May (1977) provide some analyses of the first three of these models. Getz and Haight (1989) and Hilborn et al. (1995) are more general discussions of the problems of sustainable harvesting of renewable populations.

Quota harvesting

Let’s begin by assuming that a constant amount, or quota, of N_1 is harvested by all individuals of N_2 per unit time, or that

$$f(N_1)N_2 = H \quad (7.2)$$

yielding the following model for the harvested population:

$$\frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{K} \right) - H \quad (7.3)$$

How does the subtraction of H affect the equilibria of the logistic model? The population will be at equilibrium when its growth rate (the logistic term on the r.h.s. of Eq. 7.3) equals the harvest rate H . Using a graphical technique pioneered by Noy-Meir (1975), if we plot both the logistic growth function and the total harvest function $f(N_1)N_2$ on the same graph, the equilibria correspond to the points where these two functions coincide (Fig. 7-1). Now, the rate of change of the population will be negative when the harvest function is greater than the growth function and positive when it is less. We next add arrows to our graph to indicate the direction of change of N_1 to either side of the two equilibria, N_1^{1*} and N_1^{2*} , where N_1^{1*} is the lesser and N_1^{2*} is the greater of the two equilibria for N_1 , respectively. In comparison with the equilibria of the logistic model, N_1^{1*} is displaced upward from 0 but remains unstable while N_1^{2*} is displaced downward from K but remains stable. Adding the harvest function to the logistic model has caused the two equilibria to approach one another.

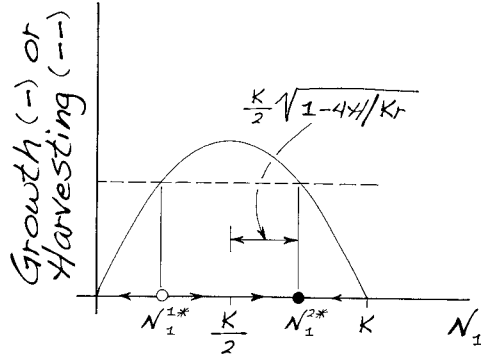
Let’s now calculate expressions for the two equilibria by setting Eq. 7.3 equal to zero and solving for N_1^* :

$$0 = rN_1 \left(1 - \frac{N_1}{K} \right) - H = -\frac{r}{K}N_1^2 + rN_1 - H$$

Using the quadratic formula (Chapter 2) to solve for N_1^{i*} we arrive at the following equilibria:

$$N_1^{i*} = \frac{1}{2}K \pm \frac{1}{2}K \sqrt{1 - \frac{4H}{Kr}} \quad (7.4)$$

Fig. 7-1 Equilibria and their stability in the quota harvesting model.



Recall from Chapter 5 that $1/2K$ is the value of N where dN/dt is maximal for the logistic model. Therefore, these new equilibria are symmetrically placed to either side of the maximal growth rate at $1/2K$ at a distance determined by the term which includes the square root sign.

Strictly and mathematically speaking, 0 is no longer an equilibrium of the population once it is harvested: when $N_1 = 0$, $dN_1/dt = -H$, not zero as an equilibrium point requires. However, biologically speaking, once N_1 is less than the unstable equilibrium point N_1^{1*} then the population density is driven to 0 and hence extinction.

Let's see if we can algebraically verify and add to our understanding of this geometric stability analysis of Eq. 7.3 by calculating eigenvalues and evaluating them at the two new equilibria. The eigenvalues of Eq. 7.3 are given by its derivative with respect to N_1 , or

$$\lambda = r - \frac{2r}{K}N_1$$

Evaluated at the two equilibria, the eigenvalues are:

$$\lambda_{N_1^{1*}} = \frac{2r}{K} \sqrt{r^2 - \frac{4r}{K}H} \quad (7.5a)$$

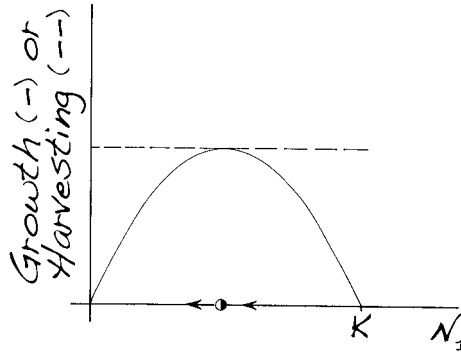
and

$$\lambda_{N_1^{2*}} = -\frac{2r}{K} \sqrt{r^2 - \frac{4r}{K}H} \quad (7.5b)$$

As in our geometric analysis, the signs of the eigenvalues show that the lesser equilibrium (N_1^{1*} , Eq. 7.5a) is unstable and the greater equilibrium (N_1^{2*} , Eq. 7.5b) is stable, so long as $H < rK/4$ (we will see the importance of this particular value of H shortly). Any population density between these two equilibria will be repelled by N_1^{1*} and converge on N_1^{2*} . Note that if the population is not harvested ($H = 0$), then we recover the eigenvalues for the logistic model (r and $-r$, respectively).

It is important to recognize that harvesting is not a perturbation in this model: the system being examined is the harvested population. A perturbation is an additional

Fig. 7-2 A saddle-node bifurcation in the quota harvesting model.

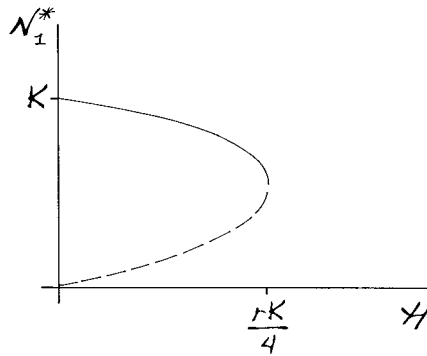


displacement away from an equilibrium due to additional mortality which occurs naturally due to poor weather or other natural disasters or augmentation of the population via immigration or stocking. How does increasing harvesting affect the response time of the model to additional perturbations? You can see that increasing H decreases the numerator of the eigenvalues and so results in slower growth away from N_1^{1*} and slower return to N_1^{2*} after a perturbation. Therefore, the population responds more sluggishly to additional perturbations near its two equilibria: increasing harvesting may result in greater yields, but slower recovery of the harvested population to unanticipated additional mortality.

It is obvious from Eq. 7.4 that as H increases to $rK/4$, the term under the square root sign approaches 0 and the two equilibria collide at $1/2K$, the maximum of the logistic growth term. At this point, from Eq. 7.5, the eigenvalue also equals 0. Is the equilibrium then stable or unstable there? To begin to answer this, plot the vectors of directions of change in N_1 when H coincides with the peak of the logistic curve (Fig. 7-2). The equilibrium is stable with respect to perturbations that move N_1 to the right (stocking) but not perturbations that move N_1 to the left (additional harvesting or other additional sources of mortality). The equilibrium is therefore a semistable equilibrium. With any further increase in H beyond $rK/4$, the equilibrium disappears entirely because now H is everywhere greater than the logistic term. Mathematically, $N_1 \rightarrow -\infty$ when $H > rK/4$, although biologically the population goes extinct ($N_1 = 0$). Therefore, this value of H is a critical point in the qualitative behavior of the model, resulting in a saddle-node bifurcation in one dimension (Fig. 7-2).

A saddle-node bifurcation occurs when two equilibria, one stable and the other unstable, approach each other, collide, and then annihilate each other with further changes in a control parameter, in this case H (Strogatz 1994). Seen from the other direction as H decreases from above towards the critical value, the equilibria appear “out of the clear blue sky,” inspiring Abraham and Shaw (1988) to call saddle-node bifurcations “blue sky bifurcations,” perhaps one of the more poetic terms in mathematics. A bifurcation diagram of the quota harvesting model of N_1^* against H shows the two branches, one stable and the other unstable, approaching one another as H increases to the critical value of $rK/4$, above which there are no equilibria at all (Fig. 7-3). Compare this diagram for a saddle-node bifurcation to Fig. 5-6, the transcritical bifurcation for the logistic model in Chapter 5, in which there are both stable and unstable branches to either side of the point where they collide.

Fig. 7-3 Saddle-node bifurcation in the quota harvesting model with increased harvesting.



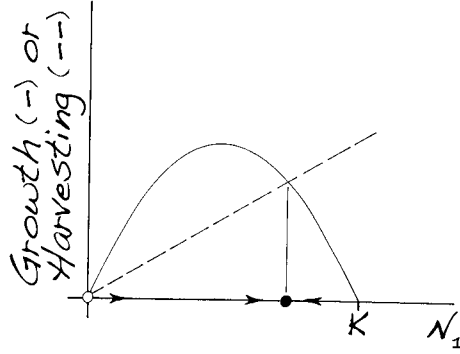
Adding a quota harvest function to the logistic model in effect translates the x -axis upward relative to the logistic model. When $H = 0$ and there is no harvest, the harvest function and the N_1 -axis lie atop one another and we recover the logistic model. The translation has caused the equilibria to approach one another and produced a new bifurcation after they collide.

It is obvious that the level of harvesting at the saddle-node bifurcation is the maximum that can be “sustained” by the population. This has led many ecologists and resource managers to call this level of harvesting the “maximum sustainable yield.” Many management plans for fisheries and timber resources are actually based on the philosophy of trying to harvest all of the maximum possible growth and sustain the population at a level equal to half its so-called “carrying capacity.” But it is obvious that we must take the term “sustainable” with a grain of salt (even a truckload of salt!). Certainly, “sustainable” in this situation does not imply “stably sustainable.” Any error in setting or implementing the harvest quota H which sets H above the critical value will result in a decrease in N_1 . If H is then maintained at this level (as the strategy of a constant quota of harvesting implies), then the population will become extinct as $N_1 \rightarrow 0$. Similar considerations apply when unforeseen circumstances cause a decline in r and K , such as a change in climate that decreases birth rates or increases death rates. Clearly, to be safe the industry or consumer must hold its quota of harvesting at some lower value than maximum. Maximizing H subject to the constraint that N_1^* exists and be positive is neither a good management nor a good evolutionary strategy!

Proportional harvesting

In the absence of precise knowledge of the population dynamics of the resource and the ability of the consumer to precisely choose the level of harvesting to remain below the saddle-node bifurcation, quota harvesting can lead to serious difficulties. Perhaps we need to relax the assumption that the harvesting rate is constant and make it depend in some way on the population density of the resource. The simplest way to do this is to assume that some proportion, h , of N_1 is harvested by each individual of N_2 , making $f(N_1) = hN_1$. This linear per capita increase in harvesting or consumption is sometimes called fixed effort harvesting because the proportion harvested is taken as a measurement of effort expended. It is often called a Holling Type I curve (Holling 1959). It is sometimes assumed that this linear increase hits a maximum where the consumer is satiated, whereupon it flattens abruptly. We will not consider this latter

Fig. 7-4 Equilibria and their stability in the proportional harvesting model at low harvesting rates.



case because the derivative at the breakpoint where the linear increase suddenly flattens is not uniquely defined, and so the function is not continuously differentiable. Such functions are difficult to analyze mathematically because one cannot say what is happening at the breakpoint. Therefore, we will assume that proportional harvesting is simply a linear increase in the per capita consumption rate of N_1 by N_2 . Inserting hN_1 for $f(N_1)$ in Eq. 7.1 we obtain:

$$\frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{K} \right) - hN_1N_2 \quad (7.6)$$

As we shall see later, the total harvesting term, hN_1N_2 , is equivalent to species interaction terms in the Lotka–Volterra equations, so proportional harvesting is also known as Lotka–Volterra interaction between two species or a species and a resource.

Using the same graphical technique as we used for quota harvesting, we see that proportional harvesting has two equilibria (Fig. 7-4). One equilibrium is obviously 0 and is unstable, just as for the logistic model. This can also be demonstrated algebraically because we can now factor out an N_1 out of all terms in Eq. 7.6, which results in $N_1^{1*} = 0$ once Eq. 7.6 is set to 0 and solved for N_1 .

What about the other equilibrium? The geometric analysis shows that this is less than K and is stable. How much less than K and how does this decrease depend on the total harvest rate? That is found by setting the remaining terms equal to 0 after factoring out an N_1 and then solving for N_1^{2*} :

$$0 = r \left(1 - \frac{N_1^{2*}}{K} \right) - hN_2$$

yielding

$$N_1^{2*} = K \left(1 - \frac{hN_2}{r} \right) = K - K \frac{hN_2}{r} \quad (7.7)$$

The second term on the r.h.s. reduces N_1^{2*} from K in proportion to hN_2 and in inverse proportion to r .

Let's examine the stability of the proportional harvest model further by evaluating the eigenvalues at the two equilibria. As always, we first calculate the derivative of Eq. 7.6 with respect to N_1 :

$$\lambda = r - \frac{2r}{K}N_1 - hN_2$$

Evaluating this at the two equilibria, the eigenvalues are:

$$\lambda_{N_1^1} = r - hN_2 \quad (7.8a)$$

and

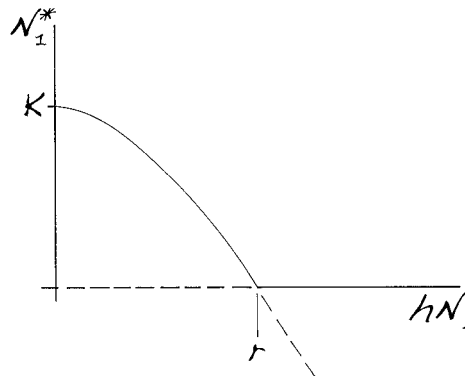
$$\lambda_{N_1^2} = hN_2 - r \quad (7.8b)$$

Note that one eigenvalue is the negation of the other, just as we have seen for the logistic model of an unharvested population. The eigenvalue at N_1^{1*} is positive when $r > hN_2$ and hence unstable. Similarly, the eigenvalue at N_1^{2*} is negative when $r > hN_2$ and hence stable. Therefore, the same conditions determine the stability of each equilibrium point.

What happens when the harvest rate, hN_2 , exceeds r ? First, when $hN_2 = r$, it can be seen from Eq. 7.7 that $N_1^{2*} = 0 = N_1^{1*}$ and the two equilibria collide. Graphically, the harvest function becomes tangent to the logistic function at $N = 0$ when $hN_2 = r$. In addition, the eigenvalues also equal 0, making the equilibrium point semistable. With further increases in hN_2 or decreases in r , the eigenvalue at N_1^{1*} becomes negative, hence $N_1^{1*} = 0$ becomes stable and the population goes extinct.

Therefore, we have a situation similar to the logistic model where the equilibria collide and exchange stabilities at a transcritical bifurcation when $r = hN_2$; the population then goes extinct. This transcritical bifurcation is clearly shown in the bifurcation diagram of N_1^* against hN_2 (Fig. 7-5). In both the logistic model and the proportional harvesting model, the same conditions determined the stability of each equilibrium point in each set: the sign of r for the logistic model and the sign of $r - hN_2$ for the proportional harvesting model. When the same conditions (or eigenvalues) determine the stabilities of two equilibrium points, only differing

Fig. 7-5 Transcritical bifurcation in the proportional harvesting model with increased harvesting.



by their sign, you should immediately suspect a transcritical bifurcation and then determine whether the two equilibria collide when these eigenvalues equal 0 and then exchange sign with further changes in the parameter values.

Much like the quota harvesting term to the logistic model translated the N_1 -axis upward, adding the proportional harvesting term to the logistic model rotates the N_1 -axis relative to the logistic model with a slope equal to hN_2 . When either h or N_2 equals zero and there is no harvest, the harvest function and the N_1 -axis lie atop one another and we recover the logistic function. The rotation has changed the position of one of the equilibrium points, but otherwise everything else remains similar.

Before we leave this discussion of the proportional harvesting model, I would like to call your attention to some of its and the quota harvesting model's more serious implications for resource use, the technology of harvest, and "sustainable" development. Although historically the quota harvesting strategy was the most common management strategy, the proportional harvest strategy is now most commonly used in Europe and North America (Rosenberg et al. 1993). The proportional harvesting model is sometimes offered as a mathematical model of the "tragedy of the commons" (Olson 1965), a concept first introduced by Lloyd (1833) then recast in a twentieth century context by Hardin (1968). As Hardin notes, the meaning of the term "tragedy" is in this context best defined by the mathematician and philosopher Alfred North Whitehead (1948) as: "the solemnity of the remorseless working of things." Whitehead goes on to say that: "this inevitableness of destiny can only be illustrated . . . by incidents which . . . involve unhappiness. For it is only by them that the futility of escape can be made evident."

How can such a simple and abstract equation illustrate the Greek concept of tragedy? Here's how. Imagine you reside in a poor village on the shore of the sea. One day, while walking along the shore, you gaze out at the sea and say to yourself: "If I had a boat and some nets, I can catch fish, sell them, and perhaps raise myself out of poverty." You do so. In effect, you increase N_2 from 0 to 1 and catch a proportion h of the fish population, N_1 . Then, another person from the village walks along the shore, sees what you are doing and, coming to the same conclusion about the road to wealth, also procures a boat and nets and begins to catch fish, thereby raising N_2 from 1 to 2. The slope of the harvest function increases and the equilibrium moves down from K to that of Eq. 7.7. But the two of you do not interfere with each other, treating the fishery as a common resource to be exploited. This continues as more and more people enter the fishery, each reasoning rationally that it is better to catch fish and lift themselves out of poverty than to sit on shore and do nothing. Each new entry further increases N_2 and decreases the equilibrium further to the left of K . Everyone is happy because they are now catching fish and generating wealth and so this continues until the harvest level reaches $\frac{1}{2}K$, the maximum growth rate of the fish population. At this point the village fishery reaches its maximum catch but the wealth of the village is not equally distributed (people who don't fish don't share in the wealth). But another villager comes along and reasons, again rationally, that it is better to catch fish than to remain in poverty. But now the growth of the population declines as the steeper slope of the harvest function intersects the logistic function of fish population growth to the left of $\frac{1}{2}K$. The fish population now begins to decline with each subsequent entry of a new person into the fishery (it is still better for each person to catch fish than not,

even though the fish population and the total catch both decline). In addition, it takes longer and longer for the fish population to recover from other sources of mortality, such as predation by other fish species or natural disasters. The fish population heads towards extinction unless one or more fishers go out of business. The same reasoning applies to the introduction of new technology that allows a single person to exploit a greater proportion of the fish population, thus increasing h as well as N_2 . Hardin (1968) makes similar arguments for the introduction of pollutants into the common resource, which has the net effect of decreasing r and K by increasing the death rate or decreasing the birth rate of fish. So long as there is no cost to the polluter of dumping toxins into the commons but there is a cost of otherwise disposing of them or retaining them, the rational short-term decision that increases an individual's wealth also decreases the ability of the common resource to supply resources for the next person.

Admittedly, this scenario is simplistic in that it and Eq. 7.6 ignores the increasing difficulty of harvesting fish as fish become scarce, thereby decreasing h . Perhaps h itself must be made a function of N_1 . But despite the simplicity of Eq. 7.6, it does help us to more precisely frame the issues which we face in our relationship to the natural world. So long as the costs of exploitation are ignored or delayed sufficiently, or so long as it is assumed that others will incur the cost, the rational short-term decision is always to enter the exploitation of the common resource, increasing N_2 and perhaps increasing h , or to dump pollution into it, decreasing r and K . The net result is to decrease population growth and slow the population's response to further perturbations, thereby remorselessly and inevitably pushing the harvested resource and "innocent bystanders" such as the by-catch of unwanted fish species to the point of extinction (Hilborn et al. 1995). This process is perhaps all too evident, as Hardin (1968) points out, in the overgrazing of public land, the overharvesting of public timber, or the overharvesting of whales and fish species on the open seas, bringing a third of all bird species and a half of all mammal species to the brink of extinction (Goombidge 1992). Globalization of world markets and the technical ability to reach and exploit any species in any place in the world's oceans have made the oceans a global common resource and have allowed for the rise of "roving bandits" – highly sophisticated fishing fleets that sequentially harvest local resources to the point of extinction and then move on in an ever-expanding spatial depletion of a target species (Berkes et al. 2006). While this is happening, these fleets continually provide an adequate supply of a harvested commodity to the world's markets and no one notices that there is a problem. Eventually, however, the market suddenly collapses when the last remaining unexploited local population is depleted. It is in this way that we are led to Whitehead's "remorseless working of things" that is the essence of tragedy and best illustrated by "incidents that involve unhappiness." One obvious way out of the dilemma is for all exploiters to submit to common regulation, but this requires the assumption of a set of common values, which is all too easily eroded by political means (Crowe 1969) and all too easy for roving bandits exploiting globally common resources to ignore as the global markets develop faster than do the international agreements and programs required to control them (Berkes et al. 2006). Unfortunately, exploiters have also become adept at modifying their behavior so that regulations are less effective (Hilborn et al. 1995). Perhaps the only way out of the tragedy of the commons is to educate people of the logical consequences of their actions and hope for the best.

Exercise 7.1

Model the increasing difficulty of exploiting N_1 as it declines by making h increase or decrease from some nominal value h_0 at $N_1 = K$:

$$h = h_0 \frac{N_1}{K}$$

Insert this new function into Eq. 7.6 in place of h and analyze the resulting model by the above methods. Does this avoid the tragedy of the commons?

Exercise 7.2

The minimum viable population is one in which the growth rate is just barely positive. Call this population size $N_{1,\text{viable}}$. From Eq. 7.6, derive an expression in terms of all the other parameters that the actual per capita harvesting rate, h , must remain below in order for the population to remain viable.

Hyperbolic or Michaelis–Menten harvesting

For both the quota and proportional harvesting functions, we have implicitly assumed that harvesting or consumption occurs continuously. But there are many situations where the intake of food or resources must be broken down into two mutually exclusive steps: the initial act of killing or biting the resource, otherwise known as “cropping,” and the various acts involved in processing the resource before it can be swallowed or otherwise utilized. For example, Spalinger and Hobbs (1992) note that many mammals cannot both bite and chew (process the food after it is bitten) at the same time because biting requires occluded incisors while chewing requires occluded molars. The nature of the mammalian jaw and dental structure precludes simultaneous occlusion of both incisors and molars (try this yourself – you can’t do it). Other examples of harvesting by two mutually exclusive steps are a logger who must stop felling trees to skid the logs and stack them on a truck and transport them to the mill for sawing or a fisher who must stop catching fish to pack them in a hold and return to port for filleting and packing. Therefore, the need to stop killing, biting, or otherwise capturing food resource in order to process what has already been taken may set an upper limit to the per capita rate of intake. A very different type of harvest function results when the act of intake or consumption must be broken into two mutually exclusive tasks of cropping and then processing the food. The following derivation is from Spalinger and Hobbs (1992) except that they use bite size instead of prey density because they are referring specifically to intake by foraging mammalian browsers or grazers. Let’s derive a per capita harvesting function for $f(N_1)$ from the following parameters and the variable N_1 :

T_c	time spent biting, felling a tree, catching fish, etc., otherwise known as “cropping” in units of t
T_p	time spent processing the food obtained during T_c in units of t
T_f	total time spent feeding or in intake = $(T_c + T_p)$ in units of t
$P_{\max}$	maximum rate of processing without pausing for a new bite, fell another tree, etc. in units of mass per t or number of prey per t per area
M_f	mass of food processed for intake during T_f in units of mass as for $P_{\max}$
h	average time spent to crop a bite, fell a tree, kill a single prey, etc. in units of t . The inverse of h is the amount cropped per unit time
N_1	prey density

First note that cropping and processing are mutually exclusive as implied by the definition of the total time spent feeding:

$$T_f = T_c + T_p$$

Intake does not happen until after the cropped food is processed (you have to chew and swallow after you take a bite). The total per capita intake rate during T_f is simply the mass of food processed during T_f or

$$I = \frac{M_f}{T_c + T_p} \quad (7.9)$$

The maximum amount of food processed for intake during T_p is:

$$M_f = P_{\max} T_p \quad (7.10)$$

The total time spent cropping as a function of time spent cropping a single bite or individual (h) and prey density (or bite size), N_1 , is:

$$T_c = \frac{hP_{\max} T_p}{N_1} \quad (7.11)$$

These two intermediate steps are the key to the rest of the derivation. Inserting both Eq. 7.10 and Eq. 7.11 into Eq. 7.9 we get:

$$I = \frac{P_{\max} T_p}{\frac{hP_{\max} T_p}{N_1} + T_p}$$

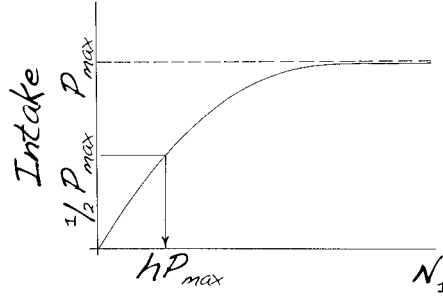
After some algebra to simplify the problem you will find that T_p cancels out both the numerator and denominator (try it!). We are then left with the following per capita harvest function:

$$I = f(N_1) = \frac{P_{\max} N_1}{hP_{\max} + N_1} \quad (7.12)$$

This function increases smoothly to an asymptotic limit $P_{\max}$, which is the maximum per capita intake rate per individual of N_2 after both cropping and subsequent processing are completed; $hP_{\max}$ is known as the half-saturation constant, or the value of N_1 where $f(N_1) = \frac{1}{2}P_{\max}$ (Fig. 7-6). It is important to note that *increases* in the half-saturation constant result in a *slower* rise to the maximum per capita intake rate.

Mathematically, functions consisting of a product of a variable and a parameter in the numerator and the sum of the variable and the same (or different) parameters in the denominator are known as hyperbolic functions. Michaelis and Menten (1913) proposed a similar functional relationship between the rate of a reaction and the concentration of the rate-limiting reactant (which serves as an analog for N_1). Their function is well known in biochemistry as the Michaelis–Menten function, although it was not until later that Briggs and Haldane (1925) presented a rigorous derivation from first principles of reaction kinetics. Monod (1958) later adapted the function

Fig. 7-6 The Michaelis–Menten function for per capita intake rate.



as a model of bacterial growth. The first use of this function in ecology was by Holling (1959), who intuitively derived it in a different manner and called it the “disk” function but it is more commonly known in population ecology as the Holling Type II function. Later, we will re-derive this same function for uptake of nutrients by plant roots. Whatever its name or derivation, this is an extremely important function throughout the whole of theoretical biology and not just ecology. An understanding of it belongs in every theoretical biologist’s toolbox. Examples of experimental verification of this function for foraging animals are given in Burnett (1956), Spalinger and Hobbs (1992), Gross et al. (1993), and Dale et al. (1994), and a modification for free-ranging moose is given in Pastor et al. (1999a).

Let’s now examine how inserting the Michaelis–Menten harvesting model for $f(N_1)$ into Eq. 7.1 affects the equilibria and their stability of a harvested population:

$$\frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{K} \right) - \frac{P_{\max} N_1}{P_{\max} h + N_1} N_2 \quad (7.13)$$

First, it is obvious that one of the equilibria is the origin because N_1 can be factored out of all terms in Eq 7.13 and the remaining N_1 in the denominator of the harvest function term can also equal zero without any difficulties. Let’s first analyze algebraically the stability of this equilibrium as the population density of the harvester, N_2 , increases to determine the conditions under which extinction of the population will happen, if at all.

The eigenvalues of Eq. 7.13 are obtained in the usual way by taking the derivative with respect to N_1 :

$$\lambda = r \left(\frac{K - 2N_1}{K} \right) - \left(\frac{P_{\max}}{P_{\max} h + N_1} - \frac{P_{\max} N_1}{(P_{\max} h + N_1)^2} \right) N_2 \quad (7.14a)$$

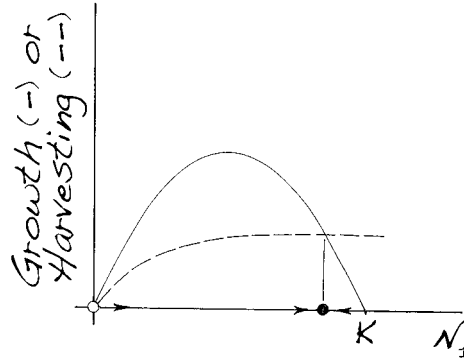
which simplifies to:

$$\lambda = r \left(\frac{K - 2N_1}{K} \right) - \left(\frac{P_{\max}^2 h}{(P_{\max} h + N_1)^2} \right) N_2 \quad (7.14b)$$

Inserting $N_1 = 0$ and evaluating, we find:

$$\lambda_{N_1^*} = r - \frac{1}{h} N_2 \quad (7.15)$$

Fig. 7-7 Equilibria and their stability in the Michaelis–Menten harvesting model.



The first term on the r.h.s is, as we have seen, the slope of the logistic growth model and the second term is the slope of the Michaelis–Menten harvesting model, both with respect to N_1 and evaluated at the origin. This eigenvalue is positive and 0 is an unstable equilibrium when $r > N_2/h$. This happens when N_2 is small, resulting in the harvesting term being less than the logistic growth term at 0. When both the harvesting and growth terms are plotted together, you can see that there is another equilibrium less than K which is also stable (Fig. 7-7).

However, as N_2 increases, the harvesting term becomes tangent to the logistic growth term at the origin and the eigenvalue at 0 then also equals zero. With further increases in N_2 , the eigenvalue at 0 becomes negative and the origin becomes stable, ensuring extinction for N_1 in the neighborhood of 0.

But something else also happens. If N_2 is not yet very large, a third equilibrium will pop into existence between the previous two as the harvesting function crosses the logistic growth function at two points in the first quadrant given by the following formidable expressions:

$$\frac{1}{2}K - \frac{hP_{\max}}{2} \pm \frac{\sqrt{r^2K^2 + 2r^2hP_{\max}K + r^2h^2P_{\max}^2 - 4rP_{\max}KN_2}}{2r} \quad (7.16)$$

Unfortunately, as models become more complex, so often do their equilibria. Inserting Eq. 7.16 into Eq. 7.14 to determine their stabilities would involve algebraic heroics which we will not do because, as we shall see in a moment, much can be inferred from simply a geometric analysis of the stabilities of these equilibria. But before we do that, note that these two equilibria are not symmetrically placed around the peak of the logistic model ($1/2K$) as in the case of the quota harvesting model. Instead, they are symmetric around a point displaced to the left of the peak by $1/2hP_{\max}$. This is because the Michaelis–Menten function asymptotically approaches the quota harvesting function and the slope of the Michaelis–Menten harvesting function near the quota harvesting asymptote is not exactly flat but “nearly” so.

It is easy to see from Fig. 7-8 that the new equilibrium point which popped into existence between 0 and the other positive equilibrium point is unstable. This unstable equilibrium between the two stable equilibria is a critical value for N_1 . If N_1 remains greater than this critical value, then it will eventually approach the greatest of the three equilibria. But if N_1 is reduced below the middle unstable equilibrium by some unforeseen circumstance, then extinction is certain.

Fig. 7-8 New equilibria and their stability in the Michaelis–Menten harvesting model as the population density of the harvester (N_2) increases.

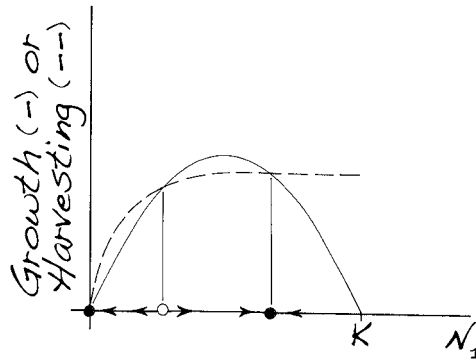
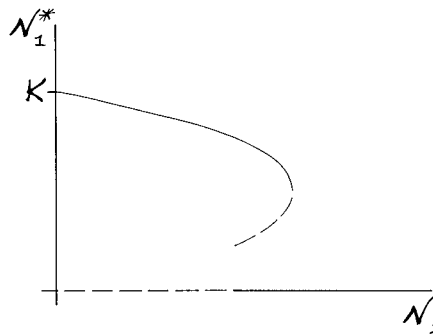


Fig. 7-9 Bifurcations in the Michaelis–Menten harvesting model as the population density of the harvester (N_2) increases.



With further increases in N_2 , the two positive equilibria approach and eventually annihilate one another in a saddle-node bifurcation, as happens with the quota harvesting model when $H = rK/4$, leaving only the origin as the sole (and now stable) equilibrium and certain extinction for N_1 . The complicated relationship amongst all three equilibria with increased N_2 is shown in Fig. 7-9.

Sigmoidal harvesting

The above harvesting functions all have one problem and that is that extinction of N_1 is certain at or above some population density or effort of the harvester, N_2 . The Michaelis–Menten harvesting function has the advantage of at least setting a maximum per capita harvesting rate because harvesting must be done in two mutually exclusive steps. Is there some way to have a maximum per capita harvest rate while at the same time avoiding certain extinction of N_1 ?

To do this, we need a harvesting function in which the derivative with respect to N_1 at $N_1 = 0$ will never be greater than the derivative of the logistic growth term. Since we are assuming that r , which is the slope of the logistic growth term at $N_1 = 0$, is always greater than 0, then having a harvesting function whose value and slope both equal 0 at $N_1 = 0$ will ensure that $N_1 = 0$ will always be unstable. Therefore, at low population densities, the harvest term will always be less than the growth term and, if there is at least one reproductive individual, the population will always grow away from 0 when it is small. Maintaining a maximum per capita harvesting rate, as in the Michaelis–Menten harvesting function, will also ensure that there is at least one stable and positive population density for N_1 (so long as the total harvesting by N_2 is maintained at some moderate level below the saddle-node bifurcation).

A little thought shows that a sigmoidal per capita harvesting function would do the trick (can you see why?). The following is a commonly used sigmoidal harvesting function:

$$f(N_1) = \frac{P_{\max} N_1^2}{h P_{\max} + N_1^2} \quad (7.17)$$

where $P_{\max}$ and h are the same parameters as for the hyperbolic (Type II) harvesting function. Eq. 7.17 is often called a Holling Type III function (Holling 1959) in the population biology literature but we will use the more descriptive term, “sigmoidal harvesting” function. It looks like the “S-curve” solution to the logistic model. (Note the squared N_1 term – if you did Exercise 7.1, where you also had a squared N_1 term, then its effect may seem vaguely familiar.)

Biologically, the decreased per capita harvesting rate as $N_1 \rightarrow 0$ could result from many factors. First, it may simply be more difficult to find food or prey when the food/prey density becomes rare. This could happen either because the probability of the harvester encountering prey becomes less and less likely, or perhaps because the harvester “looses” the search image of the prey. In addition, prey might also be able to find refuges to hide from the harvester when the prey population density is low (assuming that there are not enough refuges to hide all prey when their population density is high). For example, when prey were provided refuges which obscured their visibility when their population densities were low, the predatory wolf spider (*Pardos* sp.) switched from hyperbolic to sigmoidal harvest strategies (Döbel 1987, cited from Wise 1993, p. 156). It should be noted that there is no rigorous way to derive this function directly from more basic assumptions of cropping and processing, proportion of prey harvested per individual of N_2 , etc. (Turchin 2003). However, there is no doubt that some species employ sigmoidal per capita harvesting strategies at least some of the time (Holling 1959).

It is obvious that Eq. 7.17 meets one of our criteria of having a maximum per capita harvesting rate. Does it meet our second criterion of having a slope of 0 when $N_1 = 0$? To find out, let's take the derivative with respect to N_1 :

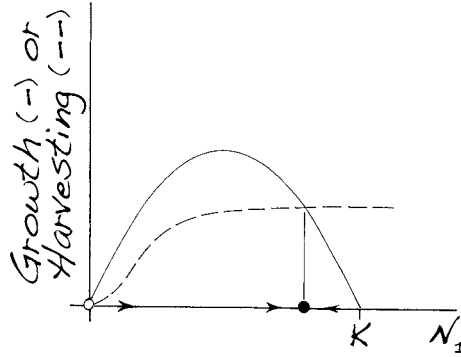
$$f'(N_1) = 2 \frac{P_{\max}^2 h N_1}{(P_{\max} h + N_1^2)^2}$$

and because the numerator of $f(N_1) = 0$ at the origin while the denominator remains positive, our second criterion is also fulfilled. Finally, inserting Eq. 7.17 into Eq. 7.1 we arrive at our model for sigmoidal harvesting of a population otherwise growing according to a logistic model:

$$\frac{dN_1}{dt} = r N_1 \left(1 - \frac{N_1}{K} \right) - \frac{P_{\max} N_1^2}{P_{\max} h + N_1^2} N_2 \quad (7.18)$$

Because N_1 can be factored out of both the logistic growth and sigmoidal harvesting terms, the origin is an equilibrium point and, because of the above considerations, it is an unstable equilibrium point. Therefore, the introduction of at least one reproductive individual of N_1 will enable the population to initially grow away from 0.

Fig. 7-10 Equilibria and their stability in the sigmoidal harvesting model at low harvesting rates.



The other algebraic forms of the equilibrium solutions of Eq. 7.18 are hopelessly unwieldy to make much sense of. Therefore, we are going to have to rely on geometric analyses of the stabilities of these other equilibria. Superimposing a graph of the sigmoidal harvesting curve atop the logistic growth curve, we find for low levels of N_2 looks the same as the corresponding graph with hyperbolic harvesting (Fig. 7-10).

But, for medium levels of N_2 (Fig. 7-11), we find something new: four equilibria alternating between unstable and stable. Now, unlike the case for Michaelis–Menten harvesting, if N_1 falls below the unstable N_1^{3*} , the population will not go extinct, but instead will be repelled from N_1^{3*} and attracted to N_1^{2*} which is greater than 0. If N_1 falls below N_1^{2*} , it will grow back to N_1^{2*} because the logistic growth term is greater than the sigmoidal harvesting term for $0 < N_1 < N_1^{2*}$. Therefore, the prey population or food resource can either go to a high or else a low equilibrium, but both are positive (Fig. 7-11). Extinction is avoided.

The alternating nature of the stable and unstable equilibria raise an important point: two nearby stable equilibria must be separated by an unstable equilibrium that demarcates their different domains of attraction. This unstable equilibrium between two stable equilibria is called a *separatrix* and we will later see generalizations of it for models of the dynamics of two or more species.

With further increases in N_2 , N_1^{4*} and N_1^{3*} collide and annihilate one another in a saddle-node bifurcation as in the quota harvesting and hyperbolic harvesting models. We are left with two equilibria: $N_1^{1*} = 0$ which is unstable and N_1^{2*} , which is stable (Fig. 7-12).

Fig. 7-11 Equilibria and their stability in the sigmoidal harvesting model at medium harvesting rates.

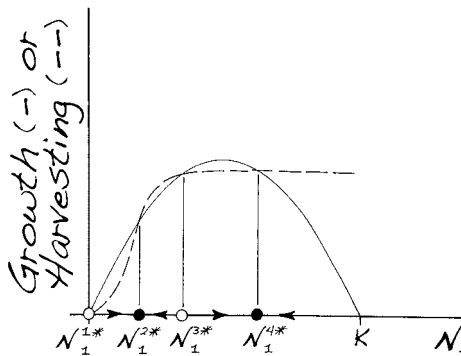
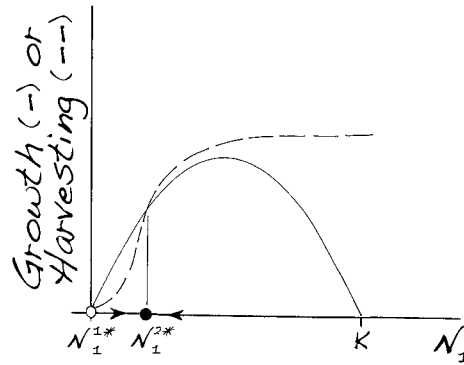


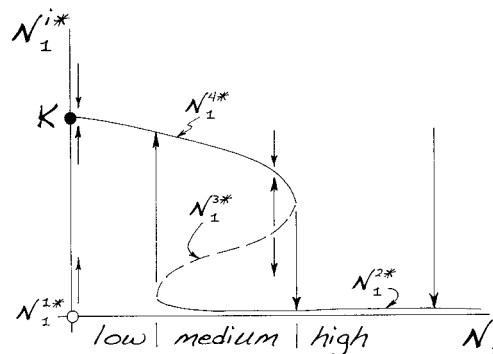
Fig. 7-12 Equilibria and their stability in the sigmoidal harvesting model at high harvesting rates.



Let's plot the bifurcation diagram (Fig. 7-13) for N_1^* against N_2 to further investigate this complicated behavior (May 1977). When N_2 is zero, there is no harvesting and $N_1^* = K$. As N_2 increases but is still low, then N_1^* declines along the upper limb of this curve. When we reach intermediate levels of N_2 , the three positive equilibria appear. At this point, besides a stable equilibrium along the upper limb, another but lower stable equilibrium also appears for the same value of N_2 , which we called N_1^{2*} . These two stable limbs are separated, and connected, by a limb of values for the unstable N_1^{3*} . This is an example of alternative stable equilibria. That is, both N_1^{1*} and N_1^{4*} are simultaneously stable at medium levels of N_2 . Which equilibrium is obtained, however, depends on the direction of change in N_2 with respect to the two threshold values which bracket the region in Fig. 7-13 labelled "medium." Because there are two stable equilibria at the same parameter values, this behavior is called *bistability*.

So long as we start with a population of N_1 at K and harvesting by increasing N_2 gradually, the equilibrium N_1 will remain on the upper limb until N_2 becomes so high that the upper limb "runs out" at the point where the saddle-node bifurcation between N_1^{3*} and N_1^{4*} happens. We are now at high levels of N_2 where N_1 jumps suddenly down to its lower stable equilibrium density N_1^{2*} . It is often said that a catastrophe has occurred at this point – prior to that we were seeing slow declines in N_1^* with increasing N_2 , but all of a sudden we have passed a threshold value of N_2 . Now, a tiny further increase in N_2 caused N_1^* to suddenly crash. Because the overall shape of the curve in this figure is that of a cusp along which a catastrophe (population crash) occurs, the behavior of a population undergoing sigmoidal harvesting is called a "cusp-catastrophe."

Fig. 7-13 Cusp-catastrophe in the bifurcation diagram for the sigmoidal harvesting model.



Now suppose we wish to reestablish N_1 at its higher equilibrium by decreasing N_2 . If we are on the lower stable limb of N_1^* and we decrease N_2 , then N_1^* will remain on the lower limb until N_2 passes completely through the intermediate level and returns to its low level. At that point, N_1 will now jump up to the higher stable limb, but not before (unless we undergo the cost of stocking it to force it up to that limb). This is an example of hysteresis – that is, the path taken by the population as one parameter (harvesting) increases is not the same as the path taken as the parameter is subsequently decreased.

The point is that resource managers who decide to employ a sigmoidal harvesting strategy should not be lulled into complacency by the fact that the harvested population cannot now be forced to extinction. It can, however, be forced suddenly to very low levels which remain stable unless harvesting is greatly curtailed and even then stocking may be necessary to bring the harvested population back to its higher stable equilibrium level within a reasonable amount of time. Once the cusp is crossed, slightly decreased harvesting does not restore the population levels. There is no free lunch here!

The cusp-catastrophe bifurcation is one of the more commonly used examples of “regime shifts” in nature (Scheffer et al. 2001, Carpenter 2003). But as we have seen, it takes a rather specialized sort of harvesting function for this bifurcation to happen. Transcritical and saddle-node bifurcations are more common in ecological models, as we shall see in future chapters.

Summary: what have we learned?

Harvesting or preying upon a single population is not a simple matter. Different harvesting strategies result in different per capita harvesting functions. When these are applied to the logistic model, they result in a rich variety of population responses and their equilibria and associated stabilities.

Small, continuous increases in harvesting result in small and continuous decreases in the harvested population – but only up to a point. Once the level of harvesting exceeds certain critical values, a small further increase in harvesting results in a discontinuous and large change in the harvested population which can be difficult to reverse, or impossible in the case of extinction. These sudden changes, or bifurcations, in the behavior of the harvested population happen at critical values of one or more of the parameters determining the magnitude of the harvest. These bifurcations can often mean extinction or, at the very least, sudden declines in the density of the harvested population. It is relatively easy to gradually reverse the effects of increased harvesting so long as the magnitude of harvesting remains below these critical levels. But once they are exceeded and the equilibria and their stability change abruptly, it can be difficult or even impossible to reverse the effects of harvesting. Such bifurcations are mathematical analogues of regime shifts in nature (Scheffer et al. 2001, Carpenter 2003).

In addition, increased harvest rates usually result in slower rates of recovery from additional perturbations near the equilibria. That is, a harvested population will recover more slowly to stable equilibria or grow more slowly away from unstable equilibria in response to additional sources of mortality compared with an unharvested population. Harvesting has a price that must be paid, and the price increases with increased harvesting!

We have encountered a variety of bifurcations in our analyses of harvested single populations, depending on the type of per capita harvesting strategy employed. Equilibria

appear, collide, disappear, and switch stabilities depending on the level and type of harvesting. Such behaviors are typical of all nonlinear models of dynamical systems such as populations. Bifurcation theory allows us to analyze where these critical points are, how they depend on parameter values, and how the different equilibria relate to one another. In this chapter, we have used both geometric and algebraic (eigenvalue analysis) approaches to exploring the behavior of equilibria and changes in their stabilities at bifurcations. The geometric analyses give us a qualitative feel for what happens, which perhaps motivates us to take up the often more difficult challenge of algebraic analysis. But the algebraic analysis gives us more precise understanding of how the equilibria and their stabilities depend on our hypothesized parameters and functional forms for population dynamics.

We can almost always analyze the behavior of the “trivial” equilibrium point $N_i^* = 0$ by calculating eigenvalues. Although mathematically “trivial” this equilibrium is not trivial biologically because it represents extinction. Therefore, the stability of this equilibrium and its dependence on parameter values should always be analyzed algebraically. However, algebraic forms of equilibria and their eigenvalues can become unwieldy for other equilibria, even in some relatively simple models. In that case, we have to rely almost solely on geometric analysis, or else “plug in” reasonable or experimentally determined values for the parameters and examine model behavior and stability for particular cases. Nonetheless, both geometric and algebraic analyses are helpful and, as we have done here, both should be used in any analysis of a nonlinear model. After all, we may get lucky and find some “nontrivial” equilibria for which it may be possible to find eigenvalues that are not too difficult to understand.

Open questions and loose ends

This chapter has presented the basic harvesting models used in the ecological literature. For further examples of other models, some of which are variations on the themes proposed here, see Turchin (2003). All in all, there seems to be a relatively complete set of harvesting functions proposed and analyzed, beginning with the logistic model as the population growth model for the harvested species. But, as far as I am aware, there has not yet been a complete analysis of how each of these harvesting functions affect the onset of period-doubling bifurcations and how they can be used to control the onset of chaos in the quadratic map or Ricker model, as we have seen in Chapter 6.

Further exploration and development of these models may come from experimental testing under controlled harvesting strategies. This can be difficult, if not impossible or even unethical, to achieve in the field. However, Fryxell et al. (2005) have recently developed experimental microcosm techniques with which they can impose different harvesting strategies on zooplankton populations. Such experimental techniques have great potential for shedding light on the behavior of harvested populations and their conformance to the models presented here. Perhaps these microscopic creatures will teach us new harvest functions with unanticipated mathematical properties.

8 Predators and their prey

Until the previous chapter, we have been assuming that a species is an ecological system unto itself. The environment surrounding a species – the other species with which it interacts and the fluxes of nonliving materials which are converted into biomass through photosynthesis and metabolism – have barely even entered our considerations. In this and the next chapter, we will relax this assumption by constructing various models of the interactions of two species. We will then generalize them to interactions amongst n -species, thereby constructing the concepts of a community or food web. Later, we will introduce constraints of mass balance of materials and energy into species interactions, thereby constructing the concept of an ecosystem. This will make the rather vague term “environmental carrying capacity” more explicit.

Perhaps one of the more basic observations in ecology is that organisms eat and are eaten. In this chapter, we will “put some flesh” on the concept of predators and their prey. What happens when we allow one species to consume another? In the previous chapter, we saw that different harvest strategies resulted in different prey equilibria with different stabilities. In this chapter we expand these ideas further by asking: How do the different harvest strategies the predator uses affect its own population dynamics as well as those of the prey it consumes?

Mathematically, this chapter will make use of the Jacobian matrix of partial derivatives and its eigenvalues (see Chapter 2). You will see that this rather abstract mathematical object is a beautiful and powerful tool with which to gauge the stability of systems of interacting parts. From this study of the Jacobian matrix of predator–prey systems will emerge a new bifurcation, the Hopf bifurcation, which spawns limit cycles between predators and their prey, perhaps one of the most studied, most fascinating, but still less-than-fully-understood aspects of ecology.

Predator–prey interactions: a first model

Before we propose a first attempt at a predator–prey model, let’s revisit the meaning of K , the so-called “environmental carrying capacity” term in the logistic equation. In Chapter 5, we derived the logistic model by assuming that the surrounding environment forced a linear decline in per capita birth rates and a linear increase in per capita death rates, without specifying the mechanism of these phenomena. We then introduced the symbol K to simplify the appearance of the logistic model. K , therefore, has nothing explicitly to do with the surrounding environment, including other species. Its connection with the environment is thus an empirical one, not one that is forced on us by the mathematics.

But, now that we are explicitly considering at least one other species interacting explicitly with the target species, we are beginning to come to grips with exactly what

the surrounding environment of a species is: in the present case, the environment consists of a predator that eats the species in question. Can we develop a model in which the consumption of the prey and the ability of the predator to convert prey biomass into its own biomass determine carrying capacity? This would allow us to do away with K for both species.

To see if predator–prey interactions alone stabilize species populations, let's dispense with logistic growth and instead assume simple exponential growth, rN_1 , for the prey, whose population density is N_1 . Let's also see if introducing a predator, whose population density is N_2 , stabilizes the exponential growth of the prey as a substitute for a carrying capacity term, K , for the prey. Now we must decide on a particular harvesting, or consumption, function for the predator, as per our considerations in Chapter 7. We assume that predation only happens when two individuals physically meet. The product of the two densities, N_1N_2 , represents the expectation that an individual predator will meet an individual prey with random movement of the two through a homogeneous landscape. The probability that, upon meeting a prey, the predator will successfully kill it will be represented by h . The total prey kill is therefore hN_1N_2 , which is, of course, the proportional harvesting model of Chapter 7. As we saw in Chapter 7, this harvesting function also has the advantage of being the simplest harvesting function which does not allow the prey species population density to become negative. Now, not all the biomass of a killed individual is transformed into predator biomass, so the growth of the predator is itself proportional to the total prey harvest, or βhN_1N_2 , where β represents the conversion efficiency of prey biomass into predator biomass, including the proportion of the kill that is consumed as well as the metabolic conversion efficiency. Finally, we assume that predators die at random with a probability m . Our first attempt at a predator–prey model then becomes:

$$\begin{cases} \frac{dN_1}{dt} = rN_1 - hN_1N_2 \\ \frac{dN_2}{dt} = \beta hN_1N_2 - mN_2 \end{cases} \quad (8.1)$$

This is the famous Lotka–Volterra predator–prey model, formulated independently by A. J. Lotka and Vito Volterra in the 1920s and 1930s. Kingsland (1995) is an excellent historical account of the work of these two pioneers of mathematical ecology. Both Lotka and Volterra were mathematical physicists, Lotka having a particular interest in physical chemistry. Both became interested in applying the mathematical techniques and rigor of physics to biology. Lotka and Volterra wanted to see if such things as mathematical laws analogous to laws in physics also could be formulated for biology. It is a telling comment that Lotka's 1925 book – which can be read even today with profit – was first entitled *Elements of Physical Biology* (later changed to *Elements of Mathematical Biology* and republished in 1956). In this book can be found, among other things, an analysis of the global carbon cycle and the effects of CO_2 on climate warming. Lotka was an American who never held a tenure position at a University, although he did some of his work as an Honorary Fellow and soft-money researcher at Johns Hopkins. The bulk of his career was spent as a physicist at General Chemical Company in New York and later analyzing population

and demographic trends with Metropolitan Life Insurance. Volterra held a chair in mathematical physics in Rome. He became interested in biology when his daughter and son-in-law, both marine biologists, asked for his help in analyzing data on Mediterranean fisheries, which led him to propose the above model. The two men never collaborated, and in fact there was some priority dispute about who had the ideas first.

We begin our analysis of this model by constructing the N_1 – N_2 phase space for this system by plotting the densities of each species and their respective equilibria on two orthogonal axes, one for the population density of each species. Any point in the phase space can serve as an initial condition for the system. Any line through the phase space is a possible trajectory through a point with the coordinates (N_1, N_2) . The trajectories taken are determined by the r.h.s. of Eq. 8.1. At each point along these trajectories there is a vector tangent to it that points in the direction in which the system is heading. The magnitude of this vector tells the rate of change of the system at that point. The vector has two components, one in the N_1 direction and one in the N_2 direction, each given by the corresponding differential equation on the r.h.s. of Eq. 8.1. The phase space is not some static space of independent points but is instead occupied by a set of vectors called a *vector field* which determines how the system gets from one point to the next. The analysis of any system of coupled differential equations is therefore an analysis of the dynamics of this vector field by geometric and algebraic means. We saw some of this in Chapter 2, but how to do this in more detail is the mathematical point of this chapter. We will go through this analysis in some detail to serve as a template for analyses of the rest of the models in this book.

Each point in phase space has one trajectory going through it and therefore can be associated with only one vector. This is a consequence of the fact that continuously differentiable equations such as Eq. 8.1 have unique solutions at each point, a unique solution guaranteeing one and only one vector at each point. If there were more than one vector at each point, then where the system headed next would be ambiguous – it could head off in either direction. Differentiability and continuity of the function or set of functions as well as its/their derivative(s) guarantees existence and uniqueness of solutions. This is called the existence and uniqueness theorem. Proof and elaboration of this theorem is a very deep and beautiful part of differential equation and dynamical system theory and is beyond the scope of this book, but you should peruse the lucid discussion of this theorem in Hirsch et al. (2004). This theorem is what allows us to make unambiguous conclusions about the behavior of these sorts of models.

The first step in analyzing the vector field is to determine the points, lines, or surfaces where the rates of change of the equations are 0: that is, where $dN_i/dt = 0$. The equilibria lie at or on these points, lines, or surfaces. In a two-species model, the equilibria for each species' population density will each lie along a line, known as a *nullcline* or, sometimes in the ecological literature, *zero-net growth isoclines* (ZNGIs) for that species. I will opt to use the shorter term *nullclines* in this book. Where the nullclines intersect, the two species will be at their equilibria simultaneously. The next step is to determine the behavior of the vector field near the equilibria by either sketching the vector field or by calculating the eigenvalues near them. This step constitutes the stability analysis of the model's equilibria. The third step is to ask how changes in parameter values affect the stability of these equilibria

– this reveals the bifurcations which separate the vector field into regions of different behaviors. As parameters cross critical values and the system goes through one of these bifurcations, it undergoes a sudden shift in its behavior which is mathematically analogous to regime shifts in natural systems (Sheffer et al. 2001, Carpenter 2003).

There are many computer programs now that will sketch the vector field for you, but knowing how to do this by hand will deepen your understanding of why it is the way it is. Let's begin our sketch of the vector field of Eq. 8.1 by solving for the nullclines. To do this, we first factor out an N_1 or an N_2 on the r.h.s. of each species differential equation in Eq. 8.1:

$$\begin{cases} f_1 = \frac{dN_1}{dt} = N_1(r - hN_2) \\ f_2 = \frac{dN_2}{dt} = N_2(\beta hN_1 - m) \end{cases} \quad (8.2)$$

We can find these nullclines by setting each of these equations equal to zero and solving for N_1 and N_2 . The first set of nullclines is of course the axes ($N_1 = 0$ and $N_2 = 0$), each axis being a nullcline for the other species. Therefore, their intersection at the origin is the trivial equilibrium. The other set of nullclines is calculated by setting the terms in parentheses in Eq. 8.4 equal to 0:

$$0 = r - hN_2$$

$$0 = \beta hN_1 - m$$

Note that each species' nullcline depends only on a critical value of the population density of the other species, not on its own population density. For example, the nullcline for N_1 is the line $N_2 = r/h$ (Fig. 8-1). When $N_2 < r/h$, losses of the prey to predation are less than the prey growth rate. Consequently, $dN_1/dt > 0$ and the prey population increases. Similarly, when $N_2 > r/h$, consumption by the predator outweighs the growth of the prey population. Consequently, $dN_1/dt < 0$ and the prey population declines. Therefore, the vector field for N_1 is a set of vectors pointing to the right below the nullcline $N_2 = r/h$ and to the left above it (Fig. 8-1).

Similarly, the nullcline for N_2 is the line $N_1 = m/\beta h$ (Fig. 8-2). When $N_1 < m/\beta h$, the capture of prey and its conversion to predators does not outpace the mortality of the predator population, and so it declines. Similarly, when $N_1 > m/\beta h$, the con-

Fig. 8-1 The nullcline for the prey (N_1) in Eq. 8.1.

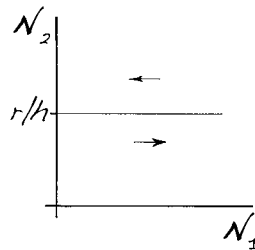


Fig. 8-2 The nullcline for the predator (N_2) in Eq. 8.1.

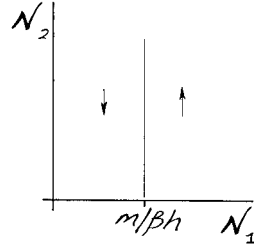
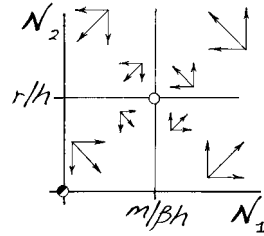


Fig. 8-3 Nullclines, equilibria and their stability, and the vector field for Eq. 8.1.



sumption and conversion of prey to predators outweighs mortality losses and the predator population increases. Therefore, the vector field for N_2 is a set of vectors pointing upward to the right of the nullcline $N_1 = m/\beta h$ and downward to the left of it (Fig. 8-2).

The vector field for the coupled system of predator and prey are the vector sums of the N_1 and N_2 components at each point (Fig. 8-3). Note that there is a counter-clockwise circulation of vectors forming a set of closed orbits concentric to the coexistence equilibrium. This is an unexpected behavior that is not apparent from a casual inspection of Eq. 8.1.

Let's examine this further by calculating the equilibria of Eq. 8.1 and the eigenvalues and eigenvectors at them. First, we solve for the equilibria by setting both equations equal to zero and solve for N_1 and N_2 simultaneously. This yields the following two equilibrium points:

$$N_1^* = 0, N_2^* = 0 \quad (8.3a)$$

$$N_1^* = \frac{m}{\beta h}, N_2^* = \frac{r}{h} \quad (8.3b)$$

The Jacobian matrix of the model is:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial N_1} = r - hN_2 & \frac{\partial f_1}{\partial N_2} = -hN_1 \\ \frac{\partial f_2}{\partial N_1} = \beta hN_2 & \frac{\partial f_2}{\partial N_2} = \beta hN_1 - m \end{bmatrix}$$

The Jacobian determines the trajectory of simultaneous changes in N_1 and N_2 and so defines the flow of the vector field in the phase plane of orthogonal axes of N_1 and N_2 . We especially wish to learn what the flow is like near equilibrium points. We therefore insert the equilibrium expressions for N_1 and N_2 (Eq. 8.3) into the elements of the Jacobian and simplify. When evaluated at the $(0, 0)$ equilibrium (Eq. 8.3a), the Jacobian becomes:

$$J|_{(0,0)} = \begin{bmatrix} r & 0 \\ 0 & -m \end{bmatrix} \quad (8.4)$$

The stability is determined by the eigenvalues, which are found by solving:

$$\begin{vmatrix} r - \lambda & 0 \\ 0 & -m - \lambda \end{vmatrix} = 0$$

which gives:

$$\lambda_1 = r, \lambda_2 = -m \quad (8.5)$$

Since one eigenvalue is positive and the other negative, the origin is an unstable saddle node.

The corresponding eigenvectors are the two axes. To see this, recall that an eigenvector, $\mathbf{w}$, associated with an eigenvalue satisfies:

$$\mathbf{J}\mathbf{w} = \lambda\mathbf{w}$$

Solving for the eigenvector associated with $\lambda_1 = r$:

$$\begin{bmatrix} r & 0 \\ 0 & -m \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = r \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}$$

which can be expressed as the following set of simultaneous equations:

$$\begin{cases} rN_1 = rN_1 \\ -mN_2 = rN_2 \end{cases}$$

These can be solved for N_2 as follows:

$$rN_1 - mN_2 = rN_1 + rN_2$$

$$0 = (r + m)N_2$$

$$N_2 = 0$$

But $N_2 = 0$ is the N_1 axis. Therefore, the N_1 component of a perturbation away from but still near the origin will grow exponentially at a rate r .

Exercise 8.1

Satisfy yourself that $N_1 = 0$, or the N_2 axis, is the eigenvector associated with $\lambda_2 = -m$. Therefore, the N_2 component of a perturbation away from but still near the origin will shrink exponentially at a rate $-m$.

By inspection of Eq. 8.1 or 8.2, without a prey ($N_1 = 0$) any predator population ($N_2 > 0$) declines at its mortality rate m to extinction, but without a predator ($N_2 = 0$), any prey population ($N_1 > 0$) grows exponentially away from 0 at a rate r . Therefore, the vector arrows point towards the origin for the N_2 axis and away from the origin for the N_1 axis.

What about the coexistence equilibrium (Eq. 8.3b)? When the Jacobian is evaluated by inserting the coexistence values for N_1^* and N_2^* found above, it becomes

$$J|_{\text{coexistence}} = \begin{bmatrix} 0 & -\frac{m}{\beta} \\ \beta r & 0 \end{bmatrix} \quad (8.6)$$

The eigenvalues are found by solving:

$$\begin{vmatrix} -\lambda & -\frac{m}{\beta} \\ \beta r & -\lambda \end{vmatrix} = 0$$

which (by taking the determinant and applying the quadratic formula) yields:

$$\lambda_j = 0 \pm i\sqrt{rm} \quad (8.7)$$

The (linearized) change of a perturbation around the coexistence equilibrium points is:

$$\begin{aligned} e^{i\lambda t} &= \cos[(\sqrt{rm})t] + i \sin[(\sqrt{rm})t] \\ e^{-i\lambda t} &= \cos[(\sqrt{rm})t] - i \sin[(\sqrt{rm})t] \end{aligned}$$

Note that the real parts of the eigenvalues are zero. Recall that the real part of the eigenvalue of a two-species model is the trace of the Jacobian, which is clearly 0 by inspection of Eq. 8.6. The determinant of Eq. 8.6 is positive ($\det(J) = 0 - \beta r(-m/\beta) = rm$). Recall also from Fig. 2-14 in Chapter 2 that when the trace of the Jacobian is 0 and the determinant is > 0 , then the equilibrium is a center and is poised at the knife-edge between stable and unstable oscillations. Such a set of equations is called *structurally unstable* because any slight change to the structure of the equations, especially changing the nonlinear terms, could tip the balance between stability or instability, depending on how changes to the structure of the equations affects the real part of the eigenvalues. Note that structural instability is not the same as a change in stability due to a bifurcation when a parameter passes through a critical value. In a bifurcation, the stability of the model changes because of a change in the numerical value of a parameter, but the form of the equations remains unchanged. In a structurally unstable equation, even slight modifications to the form of the equation alter the stability. In the remainder of this chapter, we will pose some

possible structural changes to this simplest of predator–prey models to see if we can make it stable.

The counterclockwise circulation of vectors near the coexistence equilibrium is a direct result of the purely imaginary eigenvalues of the Jacobian at that point. Any perturbation near the equilibrium causes the system to oscillate around it in perpetuity in a closed orbit, neither growing away from the equilibrium nor returning to it. The larger the perturbation, the greater the amplitude of the circulation. Furthermore, no single orbit is stable in the sense that the system converges to it after perturbations in either direction: perturb the system away from any cycle (by harvesting or stocking for example) and the system will automatically move to another cycle with different amplitude and period. Sometimes, this type of stability is called *neutral stability*, but it is really a consequence of the fact that the equilibrium is a center.

Before discussing the biological significance of this circulation, I wish to introduce two additional mathematical techniques that can sometimes help shed some light on the behavior of the vector field near these two equilibria. These techniques are both necessary but not sufficient conditions for stability and circulation of the vector field around the equilibrium point, and so conclusions using them should be made with some caution. Nonetheless, they can be useful to develop insights into the behavior of the vector field near equilibria, especially if the eigenvalues are very complicated and difficult to interpret.

The first technique is the *divergence* of a vector field $\mathbf{v}$, which is a measure of the radial departure from or convergence to an equilibrium of the vector field. The divergence of a two-dimensional vector field is calculated from elements of the Jacobian as:

$$\text{div}(\mathbf{v}) = \frac{\partial f_1}{\partial N_1} + \frac{\partial f_2}{\partial N_2}$$

evaluated at an equilibrium point. This is simply the trace of the Jacobian. You can see that the divergence of a vector field depends only on rates of change of each species population density with respect to itself, or the diagonal elements of the Jacobian. If the divergence is positive, then the vector field points away from the equilibrium point, hence it is definitely unstable. However, if the divergence is negative, the equilibrium point may be stable or it may be a saddle point (stable to perturbations along one eigenvector but unstable to perturbations along the other): recall from Fig. 2-14 that a negative trace can indicate either a semi-stable saddle point or a stable node/spiral, but a positive trace definitely indicates instability. A negative trace and divergence in a two-species model is therefore a necessary but not sufficient condition for stability. However, a positive divergence is both a necessary and sufficient condition for concluding instability. So if the trace or divergence of a vector field near an equilibrium is positive, then the equilibrium is definitely unstable.

The second mathematical technique is to calculate the *curl* of a vector field near an equilibrium, which (as its name implies) is a measure of the direction and degree of curvature of the vector field around the equilibrium. (The curl of a two-dimensional vector field is sometimes called the vorticity, the term curl sometimes being reserved for three-dimensional systems). The magnitude of the curl of a two-dimensional vector field is calculated as:

$$\text{curl}(\mathbf{v}) = \frac{\partial f_2}{\partial N_1} - \frac{\partial f_1}{\partial N_2}$$

and is evaluated at an equilibrium point. In contrast to the divergence of a vector field, the curl depends on the interaction terms, or the off-diagonal elements of the Jacobian. The curl of a two-component vector field can be used to depict a vector which points perpendicular to the plane of the vector field. The rule to determine how the curl determines the direction of circulation is known as the right hand rule. If the curl is positive, then the vector points upward. If you place your right hand on the phase plane with your thumb pointing upward, your remaining fingers curl to the left, or counterclockwise. Hence the curvature is counterclockwise if the curl is positive. Similarly, if the curl is negative, the vector points downward. Pointing the thumb of your right hand downward into the phase plane, your fingers curl to the right or clockwise.

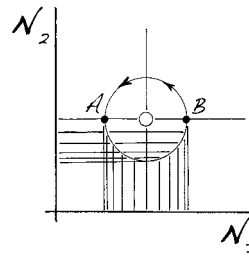
The curl indicates the average clockwise or counterclockwise circulation of a vector field, but does not necessarily imply complete circulation around an equilibrium point in closed orbits or in spirals, whereas an eigenvalue with an imaginary part does definitely imply complete circulation around the equilibrium point. In Chapter 12, we shall see an example where the curl of the vector field near a stable equilibrium point is positive and so trajectories curve into it but there are no complete circulations around the point. Therefore positive curl is a necessary but not sufficient condition for complete circulation around equilibrium points. Nonetheless, examination of the curl can be useful when eigenvalues are so complicated that it is difficult to determine the conditions when they have an imaginary part, or when eigenvalues are real but nevertheless there is still some curvature of the vector field near the equilibrium, albeit not a complete circulation.

Let's examine the divergence and curl of the vector field near each equilibrium. For the trivial equilibrium, inspection of the Jacobian (Eq. 8.6) shows that $\text{curl}(\mathbf{v}) = 0$. Therefore, there is no curvature of the vector field around this equilibrium. However, $\text{div}(\mathbf{v}) = r - m$ so the sign of the divergence depends on the relative magnitudes of r and m . If $r > m$, $\text{div}(\mathbf{v}) > 0$, the vector field points away from the origin and the system grows in the direction of larger prey densities. However if $r < m$, $\text{div}(\mathbf{v}) < 0$ and the origin is a saddle point. The predator component of the vector field points into the origin and extinction of the predator ensues. Therefore the vector field at the origin has no curl but pure divergence, the direction of which depends on the balance between the growth of the prey (r) and the death of the predators (m).

At the coexistence equilibrium, there is no divergence but pure curl, the $\text{curl}(\mathbf{v}) = \beta r + m/\beta$. Since each of these parameters is greater than 0, $\text{curl}(\mathbf{v}) > 0$ and the curvature is counterclockwise around the coexistence equilibrium, as we have seen in our sketch of the vector field.

For this simple model, calculation of divergence and curl only verifies what is easily seen in a sketch of the vector field itself. For models of three dimensions, calculation of the divergence and curl of the vector field near equilibrium points gives a check on the sketch of the vector field, but for models greater than three dimensions the curl is not defined. Further discussions of divergence and curl of a vector field and generalization to three dimensions can be found in Courant and John (1974) and Schey (2005).

Fig. 8-4 Circulation of the vector field around a counterclockwise cycle.

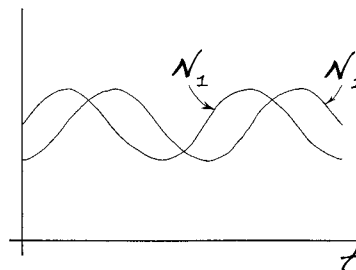


What is the biological meaning of this counterclockwise circulation around the coexistence equilibrium? If you project any closed circulation curve, or cycle, onto each axis for the linearized approximation of the model, you can see that as the coupled system goes around a circulation, the two species' population densities increase and decrease but they also lag behind one another (Fig. 8-4). For example, beginning at point A and proceeding counterclockwise to point B in equal time increments around a closed circulation path, you can see that N_1 increases continuously, at first slowly, then most rapidly when at the bottom of the closed curve, then more slowly as B is approached. By contrast, N_2 decreases, at first rapidly, then slowly at the bottom of the closed curve, then again rapidly as B is approached. Going from B to A, N_1 decreases, at first slowly, then rapidly at the top of the closed curve, then slowly again at A. At the same time, N_2 increases rapidly, then slowly at the top of the closed curve, then decreases with the decrease becoming rapid again at A.

If we plot the time series of each species population density, we get cycles that lag behind one another by $\pi/2$ radians (Fig. 8-5). Thus, the model, despite being structurally unstable, has the nice emergent property of predicting predator-prey cycles in which changes in each population drive the other with the predator's dynamics lagging behind that of the prey.

Such predator-prey cycles are well known, and their explanation has been and continues to be one of the major problems in population biology, beginning with Nicholson and Bailey (1935) and Elton and Nicholson (1942) and continuing unabated to the present (see Turchin 2003 for a recent synthesis). The emergence of predator-prey cycles in this simple first attempt at a predator-prey model is a surprising and pleasing result, and not one that could have been predicted by inspection of the original model. We now have a model that can produce coupled cycles in interacting populations (something that a single species differential equation model such as the logistic model could not do) and that at least does not diverge to

Fig. 8-5 Predator-prey cycles resulting from Eq. 8.1.



infinity when both predator and prey are present. However, there is the serious problem of structural instability of the model: to believe that the model accurately depicts nature, we must believe that nature conforms exactly to this model and not some very slight departure from it. Therefore, it seems like we are on a promising track, although admittedly we are walking a tightrope between stability and instability. Let's try to find a change to the model that hopefully will keep the cyclic behavior but also cause the cycle to be stable.

Exercise 8-2

Analyze the following model of two species that each grow according to a logistic equation but that do not interact:

$$\begin{cases} \frac{dN_1}{dt} = r_1 N_1 \left(1 - \frac{N_1}{K_1}\right) \\ \frac{dN_2}{dt} = r_2 N_2 \left(1 - \frac{N_2}{K_2}\right) \end{cases}$$

Sketch the vector field, plot the nullclines and equilibrium points and determine their stability using both geometric and algebraic methods. Compare your results to the Lotka–Volterra predator–prey model. What do adding interaction terms do to the nullclines and the vector field?

Predator–prey models: a second try with density-dependent self-regulation of prey

We wish to modify the classical Lotka–Volterra predator–prey model to make it structurally stable but still preserve the desirable property of population cycles. How should we identify the terms of the model that require changing? Since the off-diagonal elements of the Jacobian pertaining to the interaction between the species were responsible for the curl or circulation of the vector field around the coexistence equilibrium and we wish to keep this property because it causes the emergence of predator–prey cycles, we should not modify the interaction terms (at least not yet).

The instability of the model derives from the fact that predation is the only source of mortality to the prey: in the absence of any predators, the prey population grows exponentially. The obvious way to remove the instability due to exponential growth is, of course, to use a logistic equation for the first term of the prey equation. Perhaps we need to retain the use of carrying capacity after all, at least for the prey! (We are not yet finished with this business of the use of carrying capacity and will return to it when we discuss ecosystem processes such as the mass balance of nutrient cycling.)

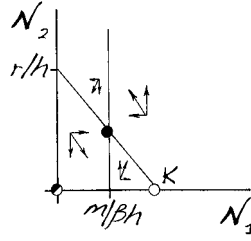
Let's then replace the exponential growth term in Eq. 8.1 with a logistic term, keeping all other terms the same, and see what happens:

$$\begin{cases} f_1 = \frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{K}\right) - hN_1N_2 = N_1 \left[r \left(1 - \frac{N_1}{K}\right) - hN_2 \right] \\ f_2 = \frac{dN_2}{dt} = \beta hN_1N_2 - mN_2 = N_2(\beta hN_1 - m) \end{cases} \quad (8.8)$$

Let's begin as before by sketching the vector field, first by finding the nullclines. One set is, of course, the axes themselves as before. Since we didn't change the predator equation, the nullcline for the predator population is the line:

$$N_1 = \frac{m}{\beta h}$$

Fig. 8-6 Nullclines, equilibria and their stability, and the vector field for Eq. 8.8 when $m/\beta h < K$.



The nullcline for N_1 is the line:

$$N_2 = \frac{r}{h} - \frac{r}{hK} N_1$$

which is a straight line with an N_2 -axis intercept r/h , a negative slope r/hK , and an N_1 -axis intercept of K . Where the nullclines intersect the two axes gives us anchor points for plotting them in the phase plane (Fig. 8-6). We see that the nullcline for the predator remains a vertical line, as before for our first predator-prey model, but the nullcline for the prey is now slanted with respect to both axes (Fig. 8-6). Note that the slanting of the nullclines with respect to one another causes the vector field to appear to spiral into the coexistence equilibrium because where the space between the two nullclines is narrowed, the N_1 component of the vector must decline rapidly to zero in order for the vector to become vertical at the N_1 nullcline.* This suggests that this equilibrium should have complex eigenvalues with negative real parts. Let's see if this conjecture holds.

First we must find the equilibria. It turns out that this model has not two but three equilibria:

$$N_1^* = 0, N_2^* = 0 \quad (8.9a)$$

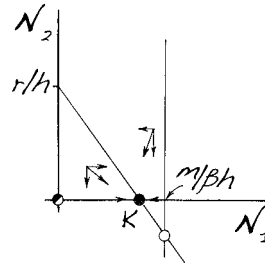
$$N_1^* = K, N_2^* = 0 \quad (8.9b)$$

$$N_1^* = \frac{m}{\beta h}, N_2^* = \frac{r}{h} \left(1 - \frac{m}{\beta h K} \right) \quad (8.9c)$$

Note that the trivial solution (Eq. 8.9a) is still an equilibrium. Furthermore, in the absence of a predator, the prey converges on K (Eq. 8.9b). This is exactly what we would predict from the logistic equation, which is what this system of equations should collapse to without a predator ($N_2 = 0$). The third, and coexistence, equilibrium (Eq. 8.9c) is new. The equilibrium population density of the prey is determined solely by parameters which pertain to the predator (h , β , and m), not on the parameter r

*If the N_1 component of the vector must be zero when the vector crosses the N_1 nullcline, the vector must be vertical. Conversely, if the N_2 component of the vector must be zero when the vector crosses the N_2 nullcline, the vector must be horizontal. Can you see why?

Fig. 8-7 Nullclines, equilibria and their stability, and the vector field for Eq. 8.8 when $m/\beta h > K$.



which pertains only to itself. The predator completely determines the equilibrium prey population through the magnitude of these parameters. By contrast, the equilibrium population density of the predator depends not only on r/h as before with exponential growth of the prey, but also on the carrying capacity of the environment for the prey. The coexistence of the predator and the prey in the third equilibrium clearly depends on some minimum value K which allows N_2^* to be positive. This sounds like a bifurcation separating the prey-only and predator-prey solutions at a critical value of K .

To find this minimum and critical K , we set the third equilibrium solution of the predator to zero and solve for K , yielding the condition $K > m/\beta h$ for coexistence. But note that $m/\beta h$ is the nullcline for the predator. This means that the nullcline for the predator must be to the left of K for coexistence: if it is to the right of K , then it intersects the nullcline for the prey in the fourth quadrant, and so the equilibrium value of the predator is negative. The coexistence solution is then unstable because if the predator population density is negative, then the predator population density decreases unboundedly; meanwhile, the consumption term hN_1N_2 is added to rather than subtracted from the logistic term in the prey equation, causing it to grow unboundedly. Conversely, the prey-alone solution is stable at K when $K < m/\beta h$ (Fig. 8-7). When $K = m/\beta h$ the two equilibrium solutions collide and exchange stability. We therefore have a transcritical bifurcation between the coexistence solution and the prey-only solution at the critical value of $K = m/\beta h$.

Let's think about this critical value a bit further. The numerator of this critical value of K is the per capita mortality of the predator; the denominator is the per capita transfer rate of prey biomass to predator biomass. This ratio therefore determines the output/input budget of the predator population. The carrying capacity of the prey population must be greater than the output/input ratio of the predator biomass in order for both to coexist. If the prey's carrying capacity is less than the output/input budget of the predator, the predator cannot survive. The prey then exists alone, and must stabilize its own population through internal density-dependent mechanisms, whatever they may be.

This critical value of K gives the minimum conditions for the smallest food chain of one link between a prey and its predator. This is the basis for the arguments of Oksanen and colleagues (Oksanen 1983, 1988, Oksanen et al. 1981) for how the length of food chains depends on the productivity of the environment, encapsulated in K . This theory says that the length of food chains shortens with decreases in environmental productivity. Eventually at low environmental productivity nothing but producers can exist stably.

Let's further examine the stability of this equilibrium point by constructing the Jacobian matrix of partial derivatives:

$$\mathbf{J} = \begin{bmatrix} r - \frac{2rN_1}{K} - hN_2 & -hN_1 \\ \beta hN_2 & \beta hN_1 - m \end{bmatrix} \quad (8.10)$$

and evaluate it at the coexistence equilibrium:

$$\mathbf{J}|_{\text{coexistence}} = \begin{bmatrix} -\frac{rm}{\beta hK} & -\frac{m}{\beta} \\ \beta r \left(1 - \frac{m}{\beta hK}\right) & 0 \end{bmatrix} \quad (8.11)$$

By inspection and a little algebra on a piece of scrap paper, you can easily convince yourself that $\text{tr } \mathbf{J}$ is negative and $\det \mathbf{J}$ is positive. Therefore, the coexistence equilibrium is stable when $K > m/\beta h$, exactly as we have seen.

To determine the behavior of the vector field near the coexistence equilibrium, we can examine in more detail the divergence, curl, and eigenvalues near it. The divergence of the vector field at this equilibrium point is also negative ($-rm/\beta hK$) and so we will now expect that the vector field has at least some components which point into the coexistence equilibrium. Again, this implies that the equilibrium may be stable. The curl of the vector field:

$$\text{curl}(\mathbf{v}) = \beta \left(r - \frac{rm}{\beta hK} \right) + \frac{m}{\beta}$$

is positive because all parameters are positive and $rm/\beta hK < r$. Therefore, the average circulation around the coexistence equilibrium is counterclockwise. But we need to examine the eigenvalues to determine if the coexistence equilibrium is stable and if the circulation results in spirals or orbits.

Solving for the eigenvalues of the Jacobian evaluated at the coexistence solution yields:

$$\lambda_j = -\frac{rm}{2\beta hK} \pm \frac{\sqrt{r^2 m^2 - 4K\beta r m(K\beta h - m)}}{2\beta hK} \quad (8.12)$$

which are complex conjugate pairs when term under the square root sign is negative. By inspection, you can see this will happen when $K > m/\beta h$, once again the value of K at the transcritical bifurcation. Since the real part of these eigenvalues is negative, there is a counterclockwise circulation of the vector field which spirals counterclockwise into the coexistence equilibrium as $t \rightarrow \infty$. Therefore, perturbations near the coexistence equilibrium die off with damped oscillations and the coexistence equilibrium is stable.

The real part of the eigenvalue gives the rate of return to equilibrium after a perturbation. The real part of the eigenvalue can be split into two parts, one containing terms pertaining to the prey (r and K) and the other containing terms pertaining to the predator (m , β , h):

$$\operatorname{Re}(\lambda_j) = -\frac{r}{K} - \frac{m}{2\beta h} \quad (8.13)$$

Since $r/K = b + d$, an increase in density-dependent self-regulation of the prey ($b + d$) increases the rate of return to equilibrium. Increasing mortality of the predator (m) also increases the rate of return to equilibrium. However, increasing per capita consumption and conversion of prey biomass into predator biomass (βh) slows the rate of return to equilibrium.

This is exactly what we saw in Chapter 7, where increased harvesting of a single species also slowed the rate of return to equilibrium. The rate of recovery to equilibrium is therefore enhanced by strong density-dependence on the part of the prey, high mortality of the predator, and a low kill-conversion efficiency of the predator. The interactions between the predator and the prey need to be weaker than the self-regulation of the prey and the mortality rate of the predator for them to coexist stably. Strong predator-prey interactions (a predator that is an efficient killer and converter of prey biomass) weaken the stability of the model. We like to think of lions, tigers, and bears as efficient killing machines, but in fact they cannot be too efficient.

In summary, both geometric and algebraic analyses confirm that at the critical value of $K = m/\beta h$, the prey-only and predator-prey solutions coincide and exchange stability as K increases further. This verifies that there is a transcritical bifurcation at this critical value of K . Indeed, this critical value and the particular bifurcation which occurs at it encapsulates the qualitative behavior of the model in the phase plane.

We now have a model that has a stable, coexistence equilibrium with predator-prey cycles, albeit damped, and so it is an improvement over the structurally unstable Lotka-Volterra model. There is, however, no evidence that predator-prey cycles in nature are damped. It is possible that the rate of damping in nature is exceedingly slow and that it therefore may not be detectable; it is also possible that episodic perturbations offset the damping and keep the system in or near a particular cycle. While plausible, these are unsatisfactory explanations: the first is by definition empirically not verifiable and the second requires us to postulate and detect these episodic events along with their magnitude and frequency distribution, making for a more complicated model. So while the new modifications to the model have been encouraging, we are not quite there yet.

A third attempt at a predator-prey model: Michaelis-Menten predator harvesting of the prey and the emergence of stable limit cycles

What should we change next? Re-invoking density-dependent growth of the prey and reinstating carrying capacity into our model stabilized the coexistence equilibrium. The mortality term for the predator seems reasonable as well. What about the harvesting term? Perhaps proportional harvesting is too simple or too constraining. Let's try a somewhat more realistic predator harvesting strategy, the hyperbolic or Michaelis-Menten harvesting function from Chapter 7, or the Type II functional response as it is sometimes called. This approach was first proposed by Rosenzweig

and MacArthur (1963), but we will develop the model and analyze its behavior somewhat differently.

Our third attempt at a predator–prey model therefore takes the following form:

$$\begin{cases} f_1 = \frac{dN_1}{dt} = rN_1\left(1 - \frac{N_1}{K}\right) - \frac{cN_1}{d + N_1}N_2 = N_1\left[r\left(1 - \frac{N_1}{K}\right) - \frac{c}{d + N_1}N_2\right] \\ f_2 = \frac{dN_2}{dt} = \beta\frac{cN_1}{d + N_1}N_2 - mN_2 = N_2\left(\beta\frac{cN_1}{d + N_1} - m\right) \end{cases} \quad (8.14)$$

For the per capita harvesting function, c is the maximum per capita “kill rate” and d is value of N_1 which gives half the per capita kill rate, or the half-saturation constant. As before, β is the metabolic transfer rate of prey biomass killed to predator biomass.

Again, there are two sets of nullclines, the axes themselves, and a second set found by equating the terms in brackets to zero and solving for N_1 and N_2 , yielding for the N_2 nullcline the vertical line

$$N_1 = \frac{md}{\beta c - m} \quad (8.15)$$

As before, the predator population increases for points that lie to the right of this nullcline and decreases for points that lie to the left. The N_1 nullcline is a parabola:

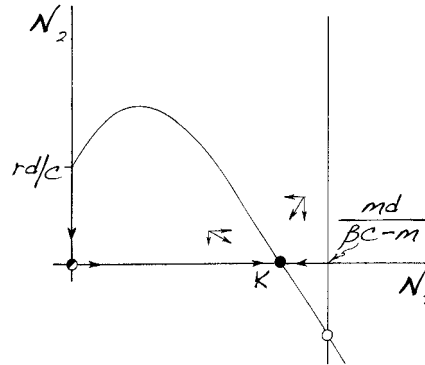
$$N_2 = \frac{rdK + r(K - d)N_1 - rN_1^2}{cK} \quad (8.16)$$

(verify this for yourself on a piece of scrap paper). As with the N_1 nullclines in the previous models, the prey population decreases for points above the parabolic nullcline and increases below it. This parabola is new – all the other nullclines we have seen so far have been straight lines. Since this parabolic shape could result in many new and interesting properties, we should explore it a bit further.

Let’s try to position it in the N_1 – N_2 phase plane by first finding its intersections with the N_1 - and N_2 -axes, which give us two anchors for it. At the N_2 -axis, $N_1 = 0$, so the intersection with the N_2 -axis happens at $N_2 = rd/c$. The intersection with the N_1 -axis happens when $N_2 = 0$ which we can find by solving the numerator using the quadratic formula, but a little inspection shows that this happens when $N_1 = K$, which is what we would expect to happen to the prey if the predator went extinct.

We can sketch the vector field by drawing both nullclines in the N_1 – N_2 phase plane anchored at their intersections with the two axes. Now something interesting happens. We are free (by choosing suitable parameter values of m , d , c , or β) to place the N_2 nullcline to the right of the N_1 nullcline altogether so that it does not intersect it in the first quadrant, or to place it so that it intersects the N_1 nullcline in the first quadrant. The intersection of the two nullclines could be to the right or to the left of the maximum of the N_1 nullcline or indeed it could be exactly at the maximum itself. All of these positions result in very different stabilities of the resulting equilibrium.

Fig. 8-8 K is stable and the predator is extinct when $md/\beta c - m > K$ for Eq. 8.14.

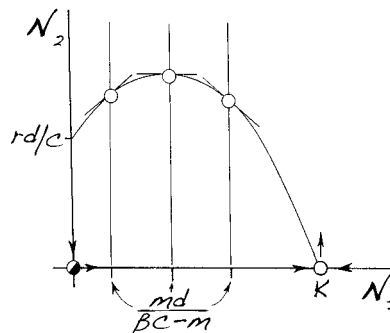


If we place the N_2 nullcline completely to the right of the N_1 nullcline in the first quadrant (i.e., to the right of $N_1 = K$), then it is easy to see that the predator goes extinct because its equilibrium value is negative (Fig. 8-8), similar to what we saw in the previous model (Fig. 8-7). Concurrently, the prey population density converges on K because the point $(K, 0)$ lies in the region of the vector field where the prey vector converges on it from both sides and the predator vector points downward. There is a stable equilibrium of $(N_1^* = K, N_2^* = 0)$ when $md/(\beta c - m) > K$ (Fig. 8-8).

If we decrease the N_2 nullcline, then the two equilibria approach each other and collide when the N_2 nullcline intersects the N_1 nullcline at K . With further decreases in the N_2 nullcline, the two equilibria exchange stabilities in a transcritical bifurcation and we have the birth of a positive, stable coexistence equilibrium (Fig. 8-9) similar to the stable coexistence equilibrium of our second predator–prey model, Eq. 8.8.

Let's decrease the N_2 nullcline further so that it approaches the maximum of the parabola from the right. To the right of this maximum, the graph is approximated by the intersection of the nullclines in our second predator–prey model (Eq. 8.8). When the N_2 nullcline intersects the parabola at its maximum, it is then perpendicular to the horizontal tangent to the parabola at that point. The graph is then approximated by the intersection of the nullclines in the classical Lotka–Volterra predator–prey model (Eq. 8.1). (Actually, Fig. 8-9 shows exactly why the classical Lotka–Volterra model is structurally unstable. Any modification to the model would move it off this point, and the two isoclines would intersect either with positive or negative slopes rather than at 90° , whereas a small modification to the

Fig. 8-9 New coexistence equilibria appear when $K > md/\beta c - m$ for Eq. 8.14.



second model would not change the topology of the intersection to the right of the maximum of the parabola.) The classical Lotka–Volterra model and our second modification of it are therefore local approximations for the situations where the nullclines intersect at the maximum of the parabolic N_1 nullcline and to the right of the maximum, respectively. It is then reasonable to expect that when the nullclines intersect at the maximum of the parabola, the nearby vector field will be closed curves as it was for the classical predator–prey model, and when the nullclines intersect to the right of the maximum, the vector field will have damped oscillations and trajectories will spiral into the coexistence equilibrium as they did for the second model.

But what happens when the nullclines intersect to the left of the maximum of the parabola? We now have a new situation that is not approximated by either of the previous models. Now the N_1 nullcline intersects the N_2 nullcline at a positive slope, unlike the negative slope at the intersection for the second model or perpendicularly as for the classical Lotka–Volterra model. If we have damped oscillations when the slope of the intersection is negative as it is to the right of the maximum, will we have undamped oscillations when the slope of the intersection is positive as it is to the left of the maximum? If so, then as the parameters m , d , β , and c are adjusted to bring the value of the N_2 nullcline to the left – that is to decrease the predator nullcline relative to the peak of the prey nullcline – we get a smooth transition from damped oscillations spiraling into a stable equilibrium point, to oscillations around a neutrally stable point, to undamped oscillations growing away from an unstable equilibrium point. Thus, we have a major qualitative change, or bifurcation, in the stability of the coexistence equilibrium as the combination of parameters for the N_2 nullcline passes through another critical value. We have all the makings of an interesting new bifurcation.

Let's find an expression for N_1 at the maximum value of its nullcline in terms of its parameters. This will give us a critical combination of parameters for N_1 against which we can judge the position of the N_2 nullcline. We can do this by taking advantage of the fact that at the peak of the nullcline, its derivative with respect to N_1 equals zero. Therefore:

$$\frac{\partial(N_1 \text{ nullcline})}{\partial N_1} = \frac{r(K - d) - 2rN_{1,\text{critical}}}{cK} = 0$$

And after cleaning up the algebra (which you should verify on a piece of scrap paper), we get

$$N_{1,\text{critical}} = \frac{1}{2}(K - d) \quad (8.17)$$

We can now set the N_2 nullcline (Eq. 8.15) equal to the r.h.s. of Eq. 8.17 to find the relationships amongst all parameters when the two nullclines intersect at the maximum of the parabola. At this point,

$$\underbrace{\frac{md}{\beta c - m}}_{N_2 \text{ nullcline}} = \underbrace{\frac{1}{2}(K - d)}_{\max N_1 \text{ nullcline}} \quad (8.18)$$

We can now solve for one parameter in terms of all the others. This will give us a critical value for that parameter that will allow us to determine where the N_2 nullcline is relative to the maximum of the N_1 nullcline. At this critical parameter value we would expect a bifurcation that qualitatively changes the stability of the coexistence equilibrium. Let's solve for K as a function of the other parameters, getting

$$K_{critical} = \frac{2md}{\beta c - m} + d \quad (8.19)$$

So now we have a critical value for one of the parameters that provides a benchmark for determining the stability of the coexistence equilibrium. When the r.h.s. is less than $K_{critical}$, we conjecture that the equilibrium point should switch stability.

The obvious next step is to calculate the equilibria and evaluate how the stability of the coexistence equilibrium changes to either side of this critical value for K . The equilibria are:

$$N_1^* = 0, N_2^* = 0 \quad (8.20a)$$

$$N_1^* = K, N_2^* = 0 \quad (8.20b)$$

$$N_1^* = \frac{md}{\beta c - m}, N_2^* = \frac{rd\beta(K\beta c - Km - md)}{K(\beta c - m)^2} \quad (8.20c)$$

At this point, calculation of the equilibria and Jacobian by hand is a somewhat heroic task. Mathematical software packages are good for this sort of thing, and more reliable than paper and pencil. Commands for calculations of equilibria, Jacobians, and eigenvalues using MatLab are given in the Appendix. Using these commands, the Jacobian matrix evaluated at the coexistence equilibrium (Eq. 8.20c) is:

$$J|_{coexistence} = \begin{bmatrix} \frac{rm[d\beta c + md + K(m - \beta c)]}{Kc\beta(\beta c - m)} - \frac{m}{\beta} & \\ \frac{r[K(\beta c - m) - md]}{K} & 0 \end{bmatrix} \quad (8.21)$$

Whether or not the equilibrium is stable depends on whether or not the trace of the Jacobian is negative and the determinant is positive. Since the trace of J is given by the sum of the diagonal elements and a_{22} (the element in the second row and column of J) is zero, the sign of the trace depends on the sign of a_{11} . Now let's substitute into this the critical value for K from Eq. 8.19. If our conjecture is correct, then the trace should be zero for this value of K , demonstrating that this critical value divides regions of the phase plane where the trace is negative (and the coexistence equilibrium may be stable) from regions where it is positive (and the coexistence equilibrium is definitely unstable). Doing this and cleaning up some messy algebra yields:

$$tr J|_{K_{critical}} = \frac{rm \cdot 0}{-2mdc\beta} = 0$$

Because the N_2 nullcline intersects the N_1 nullcline at the maximum of the parabola and simultaneously the trace of the Jacobian is zero when $K = K_{critical}$, this equilibrium is neutrally stable.

When the r.h.s. of Eq. 8.19 is less than $K_{critical}$ the N_2 nullcline then lies to the right of the maximum of the parabolic N_1 nullcline, making the trace of J negative and the determinant of J positive. The coexistence equilibrium is then stable. Because the vector field near the coexistence equilibrium is similar to that of our second predator–prey model, trajectories spiral into the coexistence equilibrium in a counterclockwise direction. On the other hand, when the r.h.s. of Eq. 8.19 is greater than $K_{critical}$ the N_2 nullcline now lies to the left of the maximum of the parabolic N_1 nullcline, making the trace of J positive. The coexistence equilibrium is now unstable and trajectories spiral away from it.

So we have verified our conjecture from the sketch of the vector field that movement of the N_2 nullcline to the left of the maximum of the N_1 nullcline parabola (which can be accomplished by increasing K) causes perturbations near the coexistence equilibrium to go smoothly from damped oscillations spiraling into a stable equilibrium, to oscillations around a neutrally stable equilibrium point, to undamped oscillations growing away from an unstable equilibrium point.

But what happens to these trajectories that spiral outward from the unstable equilibrium? Do they spiral outward without bound until they intersect one of the axes and one of the species goes extinct? Or do they settle on a particular orbit which is itself stable? Such orbits are called stable limit cycles.

To determine this, we can make use of a remarkable theorem discovered by Andrei Kolmogorov (1903–1987), one of the greatest Russian mathematicians of the twentieth century. Kolmogorov was part of the great flowering of Soviet mathematics during the 1930s and 1940s but began his mathematical career by writing a treatise on Newtonian mechanics while employed as a railroad conductor. As part of his research in dynamical system theory, Kolmogorov became interested in generalized forms of predator–prey equations. He wrote the general form of equations of two interacting populations as:

$$\begin{cases} \frac{dN_1}{dt} = N_1 f_1(N_1, N_2) \\ \frac{dN_2}{dt} = N_2 f_2(N_1, N_2) \end{cases} \quad (8.22)$$

You can see that f_1 and f_2 are the per capita growth rates for each species and are also the expressions which we have been solving for the nontrivial nullclines. Kolmogorov (1936) then proved a theorem for generalized predator–prey models stating the conditions on f_1 and f_2 for such systems to possess either a stable equilibrium point or a stable limit cycle. May (1972a) introduced this theorem to English-speaking ecologists and showed that it held for almost all forms of predator–prey models then generally in use other than the classical Lotka–Volterra model. These conditions are:

(i) the functions f_1 and f_2 are continuous and differentiable in the domain $N_1 > 0$ and $N_2 > 0$;

- (ii) $\frac{\partial f_1}{\partial N_2} < 0$
- (iii) $\frac{\partial f_1}{\partial N_1} N_1 + \frac{\partial f_1}{\partial N_2} N_2 < 0$
- (iv) $\frac{\partial f_2}{\partial N_2} \leq 0$
- (v) $\frac{\partial f_2}{\partial N_1} N_1 + \frac{\partial f_2}{\partial N_2} N_2 > 0$
- (vi) $f_1(0, 0) > 0$
- (vii) $f_1(0, A) = 0$
- (viii) $f_1(B, 0) = 0$
- (ix) $f_2(C, 0) = 0$
- (x) $B > C$

Requirements (vii) through (x) mean that there exist positive quantities A for N_2 and B and C for N_1 that satisfy them. For a proof of this theorem see Rosen (1970).

Condition (i) is satisfied by Eq. 8.14 by the existence and uniqueness theorem discussed at the beginning of this chapter. May (1975b) provides the following biological interpretations of the remaining conditions:

(ii) The per capita rate of increase of the prey decreases with increasing predator population density.

(iii) and (v) It is easiest to see the biological significance of these conditions after recognizing that (iii) is the rate of increase of the prey obtained by multiplying the population density vector $[N_1, N_2]^T$ by the first row of a Jacobian and similarly (v) is the rate of increase of the predator obtained by multiplying the population density vector by the second row. Therefore, these conditions mean that the population growth of the prey decreases but the population growth of the predator increases with increasing population densities. In other words, the more prey you have, the slower the subsequent growth rate because of self-limitation and/or predator consumption, but the faster the growth of the predator because there is more food to overcome its own mortality.

(iv) In the absence of prey, the predator population does not increase.

(vi) When both predator and prey populations are low (near zero), the prey population grows.

(vii) There is a sufficiently large predator population, $N_2 = A$, to stop the growth of the prey population when it is very small (near zero).

(viii) There is a sufficiently large prey population, $N_1 = B$, beyond which they cannot increase (due to self limitation or other limitations that prevent further growth), even if there are no predators present.

(ix) There is a minimum prey population, $N_1 = C$, that prevents further growth of the predator population.

(x) Unless $B > C$, both species will go extinct.

So the conditions have real biological meaning, as well as mathematical power. Now, taking expressions for f_1 and f_2 from Eq. 8.14, we can apply them to our model. Condition (i) is already met. For the following conditions:

$$(ii) \frac{\partial f_1}{\partial N_2} = -\frac{c}{d + N_1} < 0$$

$$(iii) \begin{aligned} N_1 \frac{\partial f_1}{\partial N_1} + N_2 \frac{\partial f_1}{\partial N_2} &= N_1 \left(\frac{cN_2}{(d + N_1)^2} - \frac{r}{K} \right) + N_2 \left(\frac{c}{d + N_1} \right) \\ &= -\frac{cd}{(d + N_1)^2} N_2 - \frac{r}{K} N_1 < 0 \end{aligned}$$

$$(iv) \frac{\partial f_2}{\partial N_2} = 0$$

$$(v) N_1 \frac{\partial f_2}{\partial N_1} + N_2 \frac{\partial f_2}{\partial N_2} = N_1 \left(\frac{\beta cd}{(d + N_1)^2} \right) + N_2 \cdot 0 > 0$$

$$(vi) f_1(0, 0) = r > 0$$

(vii) $f_1(0, A) = 0$ is satisfied if there is an $N_2 = A$ such that $r - \frac{c}{d}A = 0$, or $A = \frac{rd}{c}$. Since r, c , and $d > 0$, $A > 0$.

(viii) $f_1(B, 0) = 0$ is satisfied if there is an $N_1 = B$ such that $r \left(1 - \frac{B}{K} \right) = 0$, which of course happens when $B = K$.

(ix) $f_2(C, 0) = 0$ is satisfied if there is an $N_1 = C$ such that $\frac{\beta c C}{d + C} - m = 0$, which happens when $C = \frac{md}{\beta c - m}$. This is a particularly nice result because $\beta c > m$ for

$C > 0$. If c is the maximum per capita consumption rate of N_1 by N_2 and β is the “per kill” metabolic transfer rate from prey to predator biomass, then βc is the maximum per capita transfer rate of prey biomass to predator biomass. For $\beta c > m$, the maximum per capita transfer of biomass from predator to prey must exceed per capita predator mortality.

(x) $B > C$ is satisfied whenever $K > \frac{md}{\beta c - m}$, which is the requirement for coexistence (Fig. 8-9).

So, we are done. Every condition of Kolmogorov’s theorem is satisfied. The model contains either a stable limit cycle or a stable equilibrium. We know that there is a stable coexistence equilibrium point whenever K is between two critical values: $md/(\beta c - m)$ for the transcritical bifurcation and $(2md/(\beta c - m)) + d$ for this new bifurcation. Therefore the model must then have a stable limit cycle when K is greater than this second critical value and the equilibrium point becomes unstable. There is another important theorem, known as the Poincaré–Bendixson Theorem, which states that the unstable equilibrium point which spawns the stable limit cycle must be contained within it (in fact, Kolmogorov’s Theorem follows from the Poincaré–Bendixson Theorem). We will not go into the proof of the Poincaré–Bendixson Theorem here, but you should consult Rosen (1970), Kot (2001), or Hirsch et al. (2004) for proof and further discussion.

Exercise 8.3

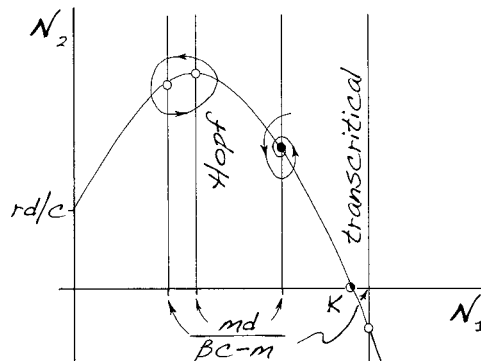
Apply Kolmogorov's theorem to the classical Lotka–Volterra (Eq. 8.1) and our second model where we modified it with a logistic equation from growth of the prey (Eq. 8.8). Does it confirm or refute our analyses of the stabilities of these models? If the model fails to satisfy Kolmogorov's Theorem, which condition or conditions are not satisfied? What is the biological meaning of not satisfying these conditions? How does this new model (Eq. 8.14) fix these problems?

There is, therefore, a bifurcation as the N_2 nullcline passes from the right to the left of the maximum of the parabolic N_1 nullcline. The stable coexistence equilibrium point then becomes unstable and is surrounded by a stable limit cycle. This new bifurcation is called a *Hopf bifurcation* (Fig. 8-10) and is closely related to the flip bifurcation in the discrete quadratic map that also spawned stable oscillations. This particular Hopf bifurcation is known as a *supercritical Hopf bifurcation* because the fixed point equilibrium becomes unstable and the vectors spiral out to a stable limit cycle. There is also something called a *subcritical Hopf bifurcation* in which an unstable limit cycle shrinks and engulfs the fixed point equilibrium, rendering it unstable and causing the vectors to spiral out to infinity (or extinction of one of the variables).

To summarize, the complete behavior of this model is determined by where the N_2 nullcline falls in relation to the N_1 parabolic nullcline (Fig. 8-10). If the N_2 nullcline lies completely to the right of the N_1 nullcline, then we have a prey species that converges to its stable carrying capacity and an unstable coexistence equilibrium in the fourth quadrant – in other words, the predator is extinct. Decreasing the N_2 nullcline causes the two equilibria to collide and exchange stabilities in a transcritical bifurcation. We now have an unstable prey-only equilibrium (K) and a new coexistence equilibrium where the nullclines intersect in the first quadrant. If they intersect to the right of the maximum of the parabolic N_1 nullcline, then the coexistence equilibrium is stable and trajectories spiral into it. If they intersect exactly at the maximum of the parabola, then the coexistence equilibrium is neutrally stable and we have a Hopf bifurcation. If they intersect to the left of the maximum of the parabola, then we have a stable limit cycle surrounding the unstable equilibrium.

The eigenvalues for the Jacobian evaluated at the coexistence solution are too complicated to be of use here unless the model is simplified considerably by a change of variables to dimensionless form or other techniques (see Kot 2001 for an

Fig. 8-10 Nullclines, equilibria and their stabilities, and bifurcations of the Rosenzweig–MacArthur model, Eq. 8.14.



excellent treatment of this model by these techniques). Unfortunately, with a change of variables the biological meaning of the parameters is often obscured. However, it is instructive to think about what must be happening to the eigenvalues. First, as we have seen for the second model, when the equilibrium is stable, the eigenvalues are complex conjugate pairs with negative real parts. As K passes through the critical value and the N_2 nullcline moves smoothly past the maximum of the N_1 nullcline, the eigenvalues must remain complex conjugate pairs or else the oscillations are lost. However, the real part must move through zero and then become positive for the trajectory to spiral away from the equilibrium point. If the model is parameterized with data, then numerical examination of critical values of parameters that cause the real part of complex conjugate eigenvalues to pass through 0 is a useful step in identifying the existence of a Hopf bifurcation.

So what keeps the trajectories from spiraling out to infinity once the Hopf bifurcation is crossed? The answer lies in the nonlinear terms, either the terms that describe the interactions between the two species or else the higher order, self-limitation terms in the logistic equation. These terms push the trajectory away from the equilibrium point towards the limit cycle when the initial displacement by the perturbation is less than the distance to the limit cycle. But at the same time, if the displacement is farther than the limit cycle, these terms also push the trajectories back into it.

Although the model has either a stable equilibrium or a stable limit cycle, in some sense passing from a stable point to a stable limit cycle is a decrease in stability because oscillations are no longer dampened. This is the basis for Rosenzweig's (1971) famous "paradox of enrichment" in which enriching the environment with a limiting nutrient and increasing the amount of biomass that can be supported (e.g., increasing carrying capacity, or K) can destabilize predator-prey systems and food webs, assuming that K is proportional to the input of a limiting nutrient. Either increasing K (Eq. 8.11) or moving the predator nullcline to the left of the maximum of the prey nullcline (Fig. 8-10) causes a Hopf bifurcation. The "paradox of enrichment" should really be called "bifurcations of enrichment."

Alternatively, from Eq. 5.4, since $K = (b_0 - d_0)/(\beta + \delta)$, you can think of a high value for K as indicating either a relatively high intrinsic growth rate, $r = b_0 - d_0$, or else a relatively weak density-dependent regulation of the prey (low value for $\beta + \delta$).

Fig. 8-11 Bifurcations caused by increasing K in the Rosenzweig-MacArthur model (Eq. 8.14): Rosenzweig's "paradox of enrichment."

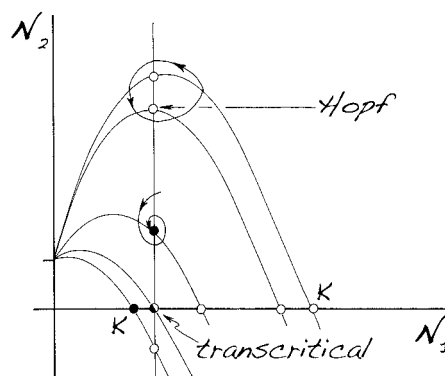
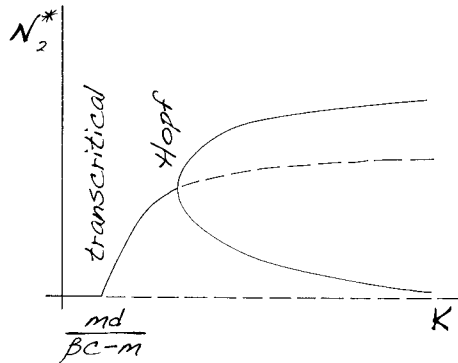


Fig. 8-12 Bifurcation diagram for the Rosenzweig–MacArthur model, Eq. 8.14, with increased K : Rosenzweig’s “paradox of enrichment.”



If weakening density-dependence of the prey destabilizes predator–prey models, then strong density-dependence of the prey would appear to be a prerequisite for stability of fixed-point equilibria in predator–prey systems. The paradox of enrichment need not necessarily result from increased nutrient input, but could also arise from weakening of density-dependent regulation by other means. A bifurcation diagram for this model (Fig. 8-12) shows both the transcritical bifurcation between the coexistence and the prey-only solutions as well as the new Hopf bifurcation as increases in K cause the real part of the eigenvalues at coexistence to pass through 0.

We are also free to vary any other parameter to move the N_2 nullcline to either side of the maximum of the parabolic N_1 nullcline. For example, increasing the mortality of the predator increases the value of the N_2 nullcline, shifting it to the right of the maximum of the parabola and therefore also stabilizing the coexistence equilibrium. Hunting the predator would accomplish this.

Unfortunately, wildlife managers often try to increase habitat quality (increase K), which destabilizes the system, at the same time as they decrease predator population density by “predator control,” which increases m and stabilizes the system. The net effect of these two opposing forces is anyone’s guess. The uncertainty of the outcome of doing two things having opposite effects is perhaps what results in unexpected surprises in wildlife management; these surprises then have the unintended effect of making both hunters and environmental groups angry. A little attention to the mathematics might help keep tempers cool and avoid some lawsuits! (More on this below.)

Exercise 8.4

Investigate the effect of additional proportional harvesting by hunters on the stability of these predator–prey models. Assume three scenarios: (i) a proportion h_1 of prey N_1 is harvested by hunters (that is subtract $h_1 N_1$ from the prey equation); (ii) a proportion h_2 of predator N_2 is harvested by hunters; (iii) both prey and predator are harvested. What do these harvesting terms do to the position of the nullclines and the stability of the equilibria? Are there critical ratios of h_1/h_2 that determine the presence and nature of stable equilibria?

Summary: what have we learned?

Relaxing the implicit assumption of single species models by introducing a second species which preys upon it results in the appearance of qualitatively new behaviors of a differential equation model, most notably oscillations around a coexistence equilibrium point. However, in order to keep the model structurally stable, the prey must still exhibit some degree of density-dependent self-limitation. This is accomplished

by having the prey grow according to a logistic growth equation (with a carrying capacity term) when the predator is absent. Logistic growth of each species in the absence of the other is necessary for any two-species model, regardless of the nature of the interactions, because otherwise any species in isolation of the other will grow exponentially.

Moreover, the growth rates of the species with respect to themselves must outweigh the strength of the interactions between the two species for the model to be stable. This can be seen by examining the formula for the determinant of the Jacobian matrix of interactions:

$$\det \mathbf{J} = \frac{\partial f_1}{\partial N_1} \frac{\partial f_2}{\partial N_2} - \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2}$$

For the model to be stable, $\det \mathbf{J}$ must be positive. Therefore, the product of the *intraspecies* partial derivatives must be greater than the product of the *interspecies* partial derivatives. Therefore, the strength of the interaction between the predator and its prey – the conversion of prey biomass into predator biomass through killing, consumption, and metabolic transformation – must not be too strong. There is a limit on how much of the prey the predator can consume and convert into its own biomass in order for it to coexist with its prey.

Furthermore, for the model to be stable the trace of the Jacobian:

$$\text{tr} \mathbf{J} = \frac{\partial f_1}{\partial N_1} + \frac{\partial f_2}{\partial N_2}$$

must be negative. Since $\partial f_1 / \partial N_1$ must be greater than zero or else the whole system collapses for lack of a self-sustaining prey base, then $\partial f_2 / \partial N_2$ must be negative and in fact outweigh the prey growth rate. In other words, the prey must have an autocatalytic term (a term that causes itself to grow in the absence of any other species) while the predator cannot be autocatalytic and in fact should not be self-sustaining in the absence of prey. These are in fact some of the biological implications of Kolmogorov's conditions for stability. A self-sustaining prey species, moderate transfer of prey biomass to predator biomass, and a predator with a relatively high mortality rate are conditions for stability of any predator–prey systems, regardless of the form of equations used. The signs of the trace and determinant of the Jacobian matrix as well as additional conditions of Kolmogorov's theorem are constraints that must be met by any stable predator–prey model, or in fact for any model of any interaction between two species as we shall see in the following chapters.

All three elaborations of the original Lotka–Volterra predator–prey model exhibit some form of coupled population oscillations. This is a surprising and pleasant result since coupled population oscillations are perhaps the most characteristic property of any predator–prey system in nature. Therefore, we need not look to environmental fluctuations as necessary causes of cycles of predators and their prey: the consumption of prey by predators alone is sufficient to produce oscillations even with a stable environment (e.g., a constant K), although a fluctuating environment will also add its own periodicities to the populations. However, these coupled population oscillations can range from unstable oscillations (the classic model) to damped

oscillations (logistic growth of prey with proportional harvesting by the predator), to a stable limit cycle. Stable limit cycles seem to be possible only when the per capita consumption of the prey by the predator is bounded by some maximum value, as with the Michaelis–Menten predator harvesting function in the Rosenzweig–MacArthur model.

Besides lending stability to coexistence solutions in the predator–prey system, the carrying capacity term, K , also determined the nature of stability of the coexistence equilibrium. This is particularly so with regard to the position of the predator nullcline. Depending on the position of the predator nullcline in relation to K and also on the shape of the function used to depict the per capita harvest of the prey by the predator, one could get a stable monoculture of the prey without the predator (prey nullcline $> K$), or coexistence with damped oscillations, neutral stability, and stable limit cycles (prey nullcline $< K$ for all three). The carrying capacity of the prey in isolation therefore plays a critical role in determining which solution is stable and the nature of its stability. For the Rosenzweig–MacArthur model with Michaelis–Menten harvesting of the prey by the predator, increases in K result in the coexistence equilibrium point becoming unstable and spawning a stable limit cycle through a Hopf bifurcation, which is a new bifurcation that we have not encountered in single-species differential equation models. If one interprets K as indicating productivity limited ultimately by nutrient input, then we get the seemingly paradoxical result that increasing nutrient input destabilizes predator–prey systems when in fact one might reasonably expect that with higher productivity there is more food for everyone and the system should be more stable, not less so (Rosenzweig 1971). But, if one interprets K as the ratio of net growth when population is low (r) to the sum of the strengths of the dependence of per capita birth and death on prey population density as in the derivation of the logistic model in Chapter 5, then increases in K could also mean a weakening of the density-dependence of the prey population. This weakening of density-dependence could happen if the prey species could somehow evolve to withstand internal social stresses that decrease density-dependence of birth and death rates (lower rates of cannibalism in the prey population, for example) and one need not necessarily invoke increased nutrient input as a cause of stable limit cycles. The phenomenon of decreased stability with increased K does not seem so paradoxical in this case. There appears to be some ambiguity in how one biologically interprets K and its effect on stability of predator–prey systems, although its mathematical definition is precise. The problems with K mentioned at the beginning of this chapter and this ambiguity of whether or not increasing K results in paradoxical behavior of two-species models is something that we will consider further in upcoming chapters. Clearly, we need to come to grips with exactly what K represents biologically.

Open questions and loose ends

Adding other predator–prey interaction forms and additional species – Further elaboration of predator–prey models produces a rich variety of additional behaviors, including other bifurcations that are not in these models. For example, Yodzis (1989) considers a predator–prey model with sigmoidal (Type III) harvesting of the prey by the predator and Freedman and Wolkowicz (1986) and Kot (2001) considers a two-species predator–prey model with “Type IV” harvesting of the prey by the predator. In this model, predator harvesting of the prey at first increases with increasing prey density, but then decreases at higher densities. Freedman and Wolkowicz (1986) and Kot (2001)

suggest that this could happen if prey are able to more effectively defend themselves as herd size increases. For example, musk oxen are more able to defend themselves against wolves when population density is great enough to form herds than when musk oxen are isolated or in small groups. Besides having all the bifurcations of the Rosenzweig–MacArthur model, the limit cycle in Kot’s model also undergoes an additional homoclinic bifurcation in which an unstable saddle node appears at one point on the limit cycle. When the predator and prey densities on the limit cycle reach this saddle node, they then “jump” off the limit cycle and head off on an unstable eigenvector. Clearly, there must be some selection pressures for preventing the appearance of a homoclinic orbit in predator–prey systems with a Type IV predator harvesting function, perhaps for all species that have herding defense behavior such as musk oxen. Perhaps there is a natural limit to emergence of more and more complicated functional forms in the responses of natural populations to each other: more complicated functional forms in species interactions do not always beget increased stability. If there are only a small number of combinations of prey growth functions and predator–prey interaction functions that produce stable and realistic population dynamics, then that would seem to imply that natural selection favors simple response functions with reliable behaviors.

Another obvious way to elaborate predator–prey models is to extend the model to three or more species. There have been some relatively recent explorations of this in the past 15 years or so (Hastings and Powell 1991, Klebanoff and Hastings 1994, McCann and Yodzis 1994, 1995). Zhang et al. (2000) analyze an interesting “two species–three stage” model in which the prey species is subdivided into a mature and immature stage and in which a predator species preys only on the immature stage with a proportional harvesting function but in addition both the immature prey stage and the predator have density-dependent self-regulation. One consequence of the Poincaré–Bendixson Theorem from which Kolmogorov’s theorem is derived is that there cannot be any more complicated behavior in two-dimensional (two-species) models than limit cycles: chaos is not possible in such models (Strogatz 1994). But in all the three-species predator–prey models cited above, chaos appears at some critical combination of parameter values, much as it was first noticed in the three-dimensional model of atmospheric turbulence studied by E. Lorenz (1996). How this emergence of a strange attractor and its properties depends on the forms of the functions used to model growth of prey population(s) in the absence of prey, the harvesting of prey by the predators(s), and the mortality of the predator(s) remains an open question.

Harvesting of predator–prey systems – Harvesting of predator–prey systems by humans is older than civilization itself; it used to be common in anthropology circles to claim that this is a defining feature of human societies. While “Man the Hunter” is no longer considered to define what it means to be human, the hunting or harvesting of predator–prey systems still proceeds and is the source of much controversy and acrimony in environmental policy, to say nothing of the source of much species extinction and near-extinction.

If you have done Exercise 8.4, you already know that the effect of harvesting predators, prey, or both on the stability of the system depends on a pretty fine balancing of parameter values and also on which functions are chosen to represent the ecology and the human harvesting policy. Lack of appreciation amongst the public and policy-makers for this fine balancing has resulted in much acrimony,

finger-pointing, and almost always a call for more data of greater precision and accuracy, as if the answer will somehow emerge from mounds of data expressed to the n^{th} decimal point with narrow standard errors of estimates. Much more headway in resolving these policy issues can be made by examining the logic of our thinking about these problems by analyzing predator–prey models with the techniques discussed in this chapter than by continued verbal argument and collection of data.

The collapse of fisheries, especially those involving whales, because of human harvesting are perhaps the most dramatic example of the instability of harvested predator–prey systems in nature. This has long been recognized, and gaining a scientific understanding of marine food webs that include whales and how exploitation affected them was the main purpose of the voyage of the *HMS Discovery* to Antarctic waters as long ago as 1925–1927 (Hardy 1967). However, it was not until the publication of May et al. (1979) that anyone attempted a mathematical analysis of the situation. By analyzing one form of a predator–prey model in which both species grow logistically but the carrying capacity of the predator is proportional to the availability of the prey, May et al. (1979) found that proportional harvesting of baleen whales by humans increased krill abundance, but proportional harvesting of krill by humans decreased abundance of baleen whales. The sensitivity of these results to various parameter values was also investigated. These results are both ecologically and mathematically interesting but were also accepted without question in a bioeconomic analysis of whaling and fisheries policy in the North Atlantic and Barents Sea (Flaaten 1988). These bioeconomic analysis and the policy conclusions drawn from it have been criticized by Yodzis (1994), who showed that conclusions drawn from the May et al. (1979) model are sensitive to which functional forms are used, and not solely to the parameter values. In particular, Yodzis showed that the form of the model influences conclusions about how predators affect and are affected by human harvesting of prey. The particular functional forms used in this model bias the policy against sustaining whale populations, regardless of how precisely the parameters for the model are determined by “mounds of data.” Indeed, in this chapter we have seen that while there are critical parameter values that determine which solution of a particular model is stable, changing the functional form of growth and/or interaction terms has a very large qualitative effect on potential behaviors, usually through the appearance of new bifurcations. It is very likely that the “collapse” of a fishery in nature corresponds to a crossing from one solution of a model to another through a bifurcation.

Since these early studies of May et al. and Yodzis, there has been a steady output of analyses of harvesting in various types of predator–prey models (see Azar et al. 1995, Zhang et al. 2000, Martin and Ruan 2001, amongst many others). Some (although not all) of these studies consider harvesting of the predator alone or the prey alone. But predators and prey are very often harvested together, either because both provide food (bears and deer), for sport (wolves and moose), or through incidental “bycatch” of industrial net fishing on the high seas. Needless to say, there doesn’t appear to be any general agreement amongst these studies, most likely because of the different number of species modeled, the different functional forms used for population growth and interactions, and the different forms used to model human harvesting policies. It is tempting to look upon these valuable studies as a collection of case studies of particular models which, taken together, points the way to a more general theory of harvesting predator–prey systems.

We do not yet know what this theory is. To develop it, we need a systematic investigation of how different growth functions, functions for how the predator harvests the prey, and functions for how humans harvest the predator, the prey, or both determine stability and bifurcations of a two-species model, to say nothing of three or more species. Obviously, such an investigation would be a large undertaking. But it would also be a comprehensive and valuable description of the stability and bifurcation behavior in the “functional space” of harvested predator–prey models. This would be of great help in developing policy, even without waiting for years for “mounds of data” on the species in question. If we could know, for example, which combination of general functional forms for both predator–prey interactions, the harvesting of both species by humans, and critical parameter values leads to extinction of one or more species, we could then rule out such scenarios as viable policies and concentrate on collecting data for better parameterization of scenarios for specific harvest policies in which stable coexistence of predators and prey is possible.

9 Competition between two species, mutualism, and species invasions

Competition between species for scarce and finite resources is a recurring theme throughout ecology. Indeed, it was the concept of competition for finite resources that motivated Darwin to develop his theories, which are some of the foundations of ecology and evolutionary biology. The key idea of competition is that a species' growth is depressed in the presence of another because each prevents the other from gaining access to some essential resource.

Mutualism, in which the two species enhance each other's growth, is the other end of the pole to competition. Even so, the basic models for the two species are remarkably similar, differing only in the signs of the interspecies interactions. The symmetry between the two interacting species within a model (which we will define in a moment) and the symmetry between competition and mutualism models will play a large role in analyzing their properties. Indeed, these symmetries unify competition and mutualism in a mathematically satisfying way.

Derivation and analysis of the simplest competition model

Although the concept of competition is plausible, whether or not competition can be detected and, if so, whether it exerts any significant control over community dynamics is a more difficult (and hotly debated) question. In part, the answer depends on exactly what is being competed for and how competition is carried out. Obviously, this is a very complex issue. Perhaps the first question to ask is: if two species were competing for some limiting but unspecified resource, what would be the net effect on the population dynamics of each other and which species properties determine the form of the dynamics? Will one species always outcompete the other and on which of its properties does this depend? Can the two species coexist? Can one species invade a monoculture of the other and, if so, under what conditions? What is the simplest model we can construct to begin to address these questions? An analysis of such a model might give us a clue as to what to observe if we suspect that two species might be competing for a resource.

Two species which are competing for the same limiting resource (nutrients, light, shelter, nesting sites, whatever) have, in the simplest formulation, a negative effect on each other proportional to their abundance. To expand the logistic model further for two competing species, we therefore simply add another term inside the parentheses summarizing the effect of each species on the other:

$$\begin{cases} \frac{dN_1}{dt} = r_1 N_1 \left(\frac{K_1 - N_1 - \alpha_{12} N_2}{K_1} \right) \\ \frac{dN_2}{dt} = r_2 N_2 \left(\frac{K_2 - N_2 - \alpha_{21} N_1}{K_2} \right) \end{cases} \quad (9.1)$$

where α_{12} indicates the per capita effect of species 2 on the abundance of species 1, and α_{21} indicates the per capita effect of species 1 on species 2; all other parameters are as before, except that subscripts indicate the particular species referred to. This is the Lotka–Volterra competition model, which is the simplest analog of competition between two species. Note that this model says nothing about what the species are competing for nor does it say anything about the mechanism of competition. It only postulates the net effect of one species on the other (a linear depression of population density or biomass in proportion to that of the other species). The model is therefore a phenomenological model of competition in that it depicts the phenomenon without explicitly identifying a mechanism. In a later chapter, we will consider the resource explicitly and we will see that some of the conclusions to be drawn shortly will have to be somewhat modified.

In the meantime, notice that there is an interesting mathematical property of this system of equations, and that is its *symmetry*. Mathematical objects are said to be symmetrical with respect to an operation if the operation does not change the appearance or behavior of the object. If we replace the subscript for every term in the equation with the subscript for the other species, the model will not be changed. Therefore, the model is symmetrical with respect to the identity of each species. But just because the original differential equations are symmetrical does not always mean that the solutions are also symmetrical. However, we shall see that in this case there is a high degree of symmetry preserved in the solutions as well. Even for the solution where one species drives the other to extinction by strong competition, there also exists a corresponding solution where the second species drives the first to extinction. Which solution is obtained – including a coexistence solution – depends on specific parameter values. In particular, the symmetry of this system of equations is sensitive to the values of α_{ij} and K_i . This should alert you to the possibility of some interesting bifurcations between solutions as α_{ij} and K_i pass through critical values. For a more in-depth treatment of symmetry and bifurcations in the solutions of differential equations, including applications in ecology and evolutionary biology, see Golubitsky and Stewart (2002).

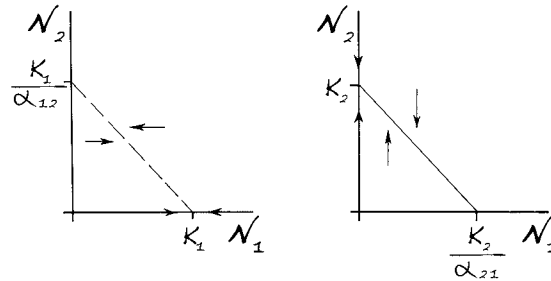
To look at the behavior and symmetries of this model further, let's find the nullclines and equilibrium solutions, then ask under what conditions the two species can coexist or not. Both the nullclines and the equilibrium solutions are found by setting either the $r_i N_i$ terms or the $(K_i - N_i - \alpha_{ij} N_j)$ terms equal 0. The nullclines for N_1 are:

$$N_1 = 0$$

$$N_1 = K_1 - \alpha_{12} N_2$$

The N_1 nullcline is reduced from K_1 in proportion to the product of N_2 and α_{12} , its per capita effect on N_1 . And similarly the nullclines for N_2 are:

Fig. 9-1 Nullclines for N_1 (left and dashed line) and N_2 (right and solid line) for the Lotka–Volterra competition model, Eq. 9.1.



$$N_2 = 0$$

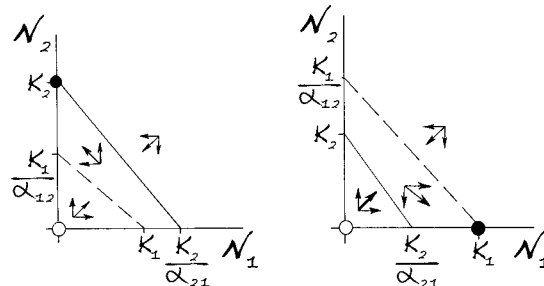
$$N_2 = K_2 - \alpha_{21}N_1$$

As with predator–prey models, the N_1 and N_2 axes are the “extinction” nullclines for the other species. If you did Exercise 8.1, you saw that with no interaction between the two species, the “nonextinction” nullclines were horizontal and vertical lines intersecting the axes at K_i . However, adding the competitive interactions causes these nullclines to “slant” with negative slopes with respect to each axis.

What happens depends on whether the relative “slant” of each nullcline causes them to intersect in the first quadrant. Each nullcline is a straight line which intersects the axes (Fig. 9-1). The N_1 nullcline intersects the N_1 axis at K_1 and intersects the N_2 axis at K_1/α_{12} . Similarly, the N_2 nullcline intersects the N_1 axis at K_2/α_{21} and the N_2 axis at K_2 . Since the vectors on each axis point to K_i , the vector components for each species point to its nullcline from left or right (for N_1) or from above or below (for N_2). The complete vector field therefore depends on whether the nullclines intersect in the first quadrant, how they intersect, and, if they do not intersect, which nullcline lies above the other.

If one species’ nullclines always lies above the other in the first quadrant (Fig. 9-2), then it is reasonable to expect that the first will always outcompete the other and drive the other to extinction. Let’s see if this is the case. Assume the N_1 nullcline always lies above the N_2 nullcline in the first quadrant. Then, both the N_1 - and N_2 -intercepts of the N_1 nullcline will be greater than the corresponding intercepts of the N_2 nullcline:

Fig. 9-2 When one nullcline lies completely above the other in the first quadrant, the equilibria are stable monocultures.



$$K_1 > \frac{K_2}{\alpha_{21}} \text{ and } K_2 < \frac{K_1}{\alpha_{12}}$$

Since ratios are harder to interpret than products, let's multiply through by the competition coefficients, yielding the following conditions:

$$\alpha_{21}K_1 > K_2 \text{ and } K_1 > \alpha_{12}K_2 \quad (9.2a)$$

The maximum competitive depression that N_1 can exert on N_2 is its per capita competitive effect on the other species (α_{21}) times its maximum biomass or population density (K_1). The maximum equilibrium biomass or population density of N_2 is K_2 since competition reduces equilibrium density from K_2 in proportion to competition from N_1 . These conditions then say that when the maximum competitive effect of N_1 on N_2 is greater than the maximum equilibrium density of N_2 and when the maximum competitive effect of N_2 on N_1 is less than the maximum equilibrium density of N_1 , then the N_1 nullcline always lies above that of N_2 in the first quadrant.

An inspection of the vector field (Fig. 9-2) shows that the trajectories that originate from points above the N_1 nullcline move the system into the region between the two nullclines or else directly to $(0, K_1)$ for points that lie close to the N_1 axis. Trajectories that originate from below the N_2 nullcline also move the system into the region between the two nullclines. Once they arrive in the region between the nullclines from either direction, trajectories move the system to $(K_1, 0)$. Therefore, it appears that $N_1^* = K_1$, $N_2^* = 0$ is a stable equilibrium point under conditions 9.2a.

Similar arguments can be made for when the N_2 nullcline always lies above the N_1 nullcline in the first quadrant, which we will call conditions 9.2b without showing it (verify this for yourself on a piece of scrap paper). We can then conclude that when the competitive depression of one species' biomass or density by another overcomes the first species' carrying capacity in isolation and when the competitive depression of the first species does not overcome the carrying capacity of the second species in isolation, then the second species will always drive the first to extinction, resulting in either one of two monocultures. Because these equilibria lie on the axes/boundaries of the first quadrant, they are called the *boundary equilibria*.

What happens if the nullclines cross? Is there always a stable coexistence equilibrium point at the intersection? The answer is that it depends on how they cross. If they cross such that

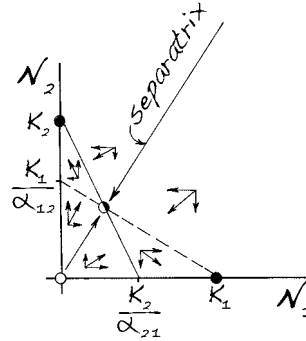
$$K_2 > \frac{K_1}{\alpha_{12}} \text{ and } K_1 > \frac{K_2}{\alpha_{21}}$$

or

$$\alpha_{12}K_2 > K_1 \text{ and } \alpha_{21}K_1 > K_2 \quad (9.2c)$$

then both species are "strong" competitors against the other. An examination of the vector field (Fig. 9-3) shows that either species can displace the other to extinction, depending on the initial conditions. Above or below both nullclines, the vectors move the system into the two triangular regions between the nullclines (except of course

Fig. 9-3 Strong competition results in stable monocultures of either species.



for points near the two axes, where the trajectories will terminate at the nearest K_i . Once in the triangular regions, the trajectories will move the system to the monoculture equilibrium point on that side of the intersection. It is apparent that the species which is initially most abundant will always displace the other. The intersection of the nullclines is therefore a saddle-node: it will be approached only if the system initially lies on the line dividing the two regions of local stability of $(K_1, 0)$ and $(0, K_2)$. This line is called a *separatrix* since it separates the two regions. Except for perturbations exactly on the separatrix, any perturbation away from the coexistence equilibrium will be attracted to the nearest monoculture, and hence boundary, equilibrium. This is another example of bistability which we first saw in the sigmoidal harvesting model in Chapter 7, or two alternative stable equilibria under the same parameter values.

Is there no hope of stable coexistence between competitors? Are the myriad examples of coexistence we see around us simply transient states on their way to monocultures? No – stable coexistence is apparently possible only if both species are weak competitors as defined by the following conditions:

$$K_2 < \frac{K_1}{\alpha_{12}} \text{ and } K_1 < \frac{K_2}{\alpha_{21}}$$

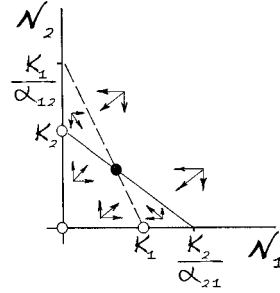
or

$$\alpha_{12}K_2 < K_1 \text{ and } \alpha_{21}K_1 < K_2 \quad (9.2d)$$

The vector field (Fig. 9-4) now shows that once trajectories arrive in the triangular regions between the two nullclines, the vectors point to the coexistence equilibrium, which appears to be stable to perturbations away from it in any direction. Because the coexistence equilibrium is in the interior of the first quadrant, it is often called the *interior equilibrium*.

Thus, stable coexistence between two competitors requires relatively weak competition between the two species. This is reminiscent of the findings of Chapter 8 that stable coexistence between a predator and its prey requires relatively weak predation. Moreover, note the determining role played by K_i in all these equilibrium. Recall that K is determined by the strength of density-dependent self-regulation of

Fig. 9-4 Weak competition results in stable coexistence.



a population. For coexistence to be stable, it appears that competition between each species cannot be stronger than self-regulation of populations in the absence of other species.

These geometric solutions appear robust, but let's see if we can verify them algebraically by calculating the equilibrium points and examining the stability of the Jacobian at each. The equilibrium points are:

$$N_1^* = N_2^* = 0 \quad (9.3a)$$

$$N_1^* = K_1, N_2^* = 0 \quad (9.3b)$$

$$N_1^* = 0, N_2^* = K_2 \quad (9.3c)$$

$$N_1^* = \frac{\alpha_{12}K_2 - K_1}{\alpha_{12}\alpha_{21} - 1}, N_2^* = \frac{\alpha_{21}K_1 - K_2}{\alpha_{12}\alpha_{21} - 1} \quad (9.3d)$$

Equilibrium 9.3a is where both species go extinct (the so-called trivial equilibrium). Equilibria 9.3b and 9.3c are where one species goes extinct and the other reaches its carrying capacity (the two monoculture equilibria). Equilibrium 9.3d is the only coexistence equilibrium. This equilibrium occurs, of course, at the intersection of the two nullclines specified by conditions 9.2d. Note the symmetries in all these solutions: if you switch N_1 and N_2 and every similarly labeled parameter, you recover the solution for the other species.

The Jacobian matrix in general is:

$$J = \begin{bmatrix} \frac{r_1(K_1 - 2N_1 - \alpha_{12}N_2)}{K_1} & -\frac{r_1\alpha_{12}N_1}{K_1} \\ -\frac{r_2\alpha_{21}N_2}{K_2} & \frac{r_2(K_2 - 2N_2 - \alpha_{21}N_1)}{K_2} \end{bmatrix} \quad (9.4)$$

The symmetry of the Jacobian is readily apparent and expected from taking the partial derivatives of Eq. 9.1. Let's evaluate this at the two monoculture solutions and the coexistence solution and see if we can develop any insights into when coexistence or monocultures develop. The Jacobian evaluated at the N_1 monoculture equilibrium is:

$$J|_{(K_1,0)} = \begin{bmatrix} -r_1 & -r_1\alpha_{12} \\ 0 & \frac{r_2(K_2 - \alpha_{21}K_1)}{K_2} \end{bmatrix} \quad (9.5)$$

for which the eigenvalues are:

$$\lambda_1 = -r_1 \text{ and } \lambda_2 = \frac{r_2(K_2 - \alpha_{21}K_1)}{K_2} \quad (9.6)$$

The first eigenvalue is already negative. The second is negative only when $\alpha_{21}K_1 > K_2$ or $\alpha_{21} > K_2/K_1$. As we have seen from our geometric analyses, this is one of the conditions for a stable monoculture of N_1 (i.e., $(K_1, 0)$). When this inequality is switched, then the N_1 monoculture becomes unstable. We can see the beginnings of a critical value of α_{21} leading to bifurcation here.

As you might expect by now, we get similar conditions for a stable monoculture of N_2 . The Jacobian evaluated at the N_2 monoculture equilibrium is:

$$J|_{(0,K_2)} = \begin{bmatrix} \frac{r_1(K_1 - \alpha_{12}K_2)}{K_1} & 0 \\ -r_2\alpha_{21} & -r_2 \end{bmatrix} \quad (9.7)$$

for which the eigenvalues are:

$$\lambda_1 = -r_2 \text{ and } \lambda_2 = \frac{r_1(K_1 - \alpha_{12}K_2)}{K_1} \quad (9.8)$$

By a similar reasoning as above for a stable N_1 monoculture, a stable monoculture of N_2 happens when $\alpha_{12}K_2 > K_1$, or when $\alpha_{12} > K_1/K_2$. This also confirms our geometric stability analysis.

What happens when $\alpha_{ij} < K_i/K_j$? When α_{ij} is less than K_i/K_j , not only will both monocultures become unstable, but the coexistence equilibrium becomes stable. $\alpha_{ij} = K_i/K_j$ is therefore a critical value for a bifurcation separating monocultures from coexistence of competitors. To see why, let's evaluate the Jacobian at the coexistence equilibrium:

$$J|_{\text{coexistence}} = \begin{bmatrix} \frac{r_1(K_1 - \alpha_{12}K_2)}{K_1(\alpha_{12}\alpha_{21} - 1)} & \frac{r_1\alpha_{12}(K_1 - \alpha_{12}K_2)}{K_1(\alpha_{12}\alpha_{21} - 1)} \\ \frac{r_2\alpha_{21}(K_2 - \alpha_{21}K_1)}{K_2(\alpha_{12}\alpha_{21} - 1)} & \frac{r_2(K_2 - \alpha_{21}K_1)}{K_2(\alpha_{12}\alpha_{21} - 1)} \end{bmatrix} \quad (9.9)$$

The eigenvalues for this are unwieldy, so let's evaluate its stability by asking: what are the conditions for making the trace negative and the determinant positive? Although each of the terms appears complicated, the answers are actually quite simple. The trace will be negative if $\alpha_{ij} < K_i/K_j$ and the determinant will be positive

if $\alpha_{21}\alpha_{12} < 1$ (verify these conditions for yourself). The conditions for stable coexistence we derived geometrically will then also be met.

To summarize, if the two nullclines cross and if $\alpha_{21} > K_2/K_1$ and $\alpha_{12} > K_1/K_2$, then we have two stable monocultures (which one is reached depends on initial conditions) separated by an unstable coexistence equilibrium. Reducing both competition parameters is equivalent to weakening the competitive abilities of both species. This rotates the nullclines toward each other around a pivot point which is the coexistence equilibrium (see Figs 9-3 and 9-4 to see why). The bifurcation that occurs at $\alpha_{ij} = K_i/K_j$ is a transcritical bifurcation because the monoculture solutions have switched stabilities simultaneously with the coexistence solution. Therefore, these α_{ij} are critical parameter values separating strong competitors (leading to stable monocultures) from weak competitors (which can coexist stably) in the simplest Lotka–Volterra competition model.

Are these results specific only to this very simple (perhaps overly simple) competition model? Ayala et al. (1973) analyzed 10 different variations on the theme of this model, including ones with more parameters and other functional forms of competition. Each of the 10 models examined also assumed logistic growth in the absence of competition, and the equations for each species were symmetric with respect to exchanging variables. In all cases, stable coexistence was possible only when $\alpha_{21}\alpha_{12} < 1$ and when a line joining the two K_i intercepts of the nullclines with the axes remained below the intersection of the two nullclines, exactly the same conditions required here. Therefore, the conditions for stable coexistence may be generalizable to any completely symmetric competition model with logistic growth forms of each species in isolation if it meets these conditions.

What about stable limit cycles? Inspection of Fig. 9-4 suggests that there cannot be coexistence by means of limit cycles in these two-species competition models. Once the trajectory enters one of the triangular regions between the nullclines, it is attracted to the equilibrium point. For a limit cycle to be possible, the trajectory must be able to exit the triangular regions, but this seems to be prohibited by the structure of the vector field. Hence, limit cycles do not appear possible.

We can prove this algebraically by recalling that the eigenvalues of a 2×2 matrix can be calculated using the following formula from Chapter 2:

$$\lambda_i = \frac{tr \pm \sqrt{tr^2 - 4\det}}{2}$$

In order for stable limit cycles to happen, eigenvalues must be complex with positive real parts. In terms of the above formula, the trace of the Jacobian must be positive and the term under the square root sign, known as the discriminant, must be negative. The square of the trace is:

$$tr^2 = \left(\frac{\partial f_1}{\partial N_1} + \frac{\partial f_2}{\partial N_2} \right)^2 = \frac{\partial f_1^2}{\partial N_1} + 2 \frac{\partial f_1}{\partial N_1} \frac{\partial f_2}{\partial N_2} + \frac{\partial f_2^2}{\partial N_2}$$

And four times the determinant is:

$$4\det = 4 \frac{\partial f_1}{\partial N_1} \frac{\partial f_2}{\partial N_2} - 4 \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2}$$

Subtracting the two, we obtain the discriminant:

$$\begin{aligned} \text{discriminant} &= \left(\frac{\partial f_1}{\partial N_1}^2 - 2 \frac{\partial f_1}{\partial N_1} \frac{\partial f_2}{\partial N_2} + \frac{\partial f_2}{\partial N_2}^2 \right) + 4 \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2} \\ &= \left(\frac{\partial f_1}{\partial N_1} - \frac{\partial f_2}{\partial N_2} \right)^2 + 4 \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2} \end{aligned}$$

Since the first term is squared, it is clearly positive. The sign of the discriminant, and therefore whether the eigenvalue is real or complex, therefore depends on the sign of the term with the products of the species interactions, $4 \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2}$. The signs of the Jacobian matrix for two-species competition models is:

$$\begin{bmatrix} + & - \\ - & + \end{bmatrix}$$

Both of the off-diagonal, species interaction terms are negative because they depict the reduction in growth of each species because of competition from the other. The product of the species interactions is therefore positive. Adding this to the positive value for the trace squared results in a discriminant which must be positive. Hence, complex eigenvalues and therefore stable limit cycles are not possible for two-species competition models.

As a check on this proof, let's see if it predicts the possibility of a stable limit cycle for predator–prey models. The signs of the Jacobian matrix for two-species predator–prey models is:

$$\begin{bmatrix} + & - \\ + & 0, - \end{bmatrix}$$

The signs of the diagonal terms are required by conditions (ii) and (iv) of Kolmogorov's Theorem (see Chapter 8). Because the off-diagonal, species interaction terms are of opposite signs, their product in the discriminant is negative. Complex eigenvalues are therefore possible in predator–prey models if $tr^2 < 4\det$. A Hopf bifurcation, and therefore the possibility of stable limit cycles, will happen if $tr > 0$.

There is an elegant experimental design to determine whether two species reduces each growth linearly (the core assumption of this model) and also to parameterize the magnitudes of the competition coefficients, α_{ij} . This is known as a DeWit replacement series (see Harper (1979) for excellent examples for plants and Ayala (1972) for examples using *Drosophila*). The experiment consists of establishing blocks or cultures of two species in various initial proportions, from 100% N_1 to 100% N_2 with various ratios such as 75% N_1 :25% N_2 , 50% N_1 :50% N_2 , 25% N_1 :75% N_2 or any other appropriate ratio, in between. The two extremes represent, of course, the monoculture biomass of magnitude K_i under the particular environmental conditions of the experiment. When the yield of each species is plotted against the other, it can be determined whether increased abundance of each species causes the abundance of the other to decrease linearly (decrease of one proportional to the abundance of the other). If so, then the slopes of the decreases are simply the

appropriate value for the competition coefficients. If the decline is not linear, then Eq. 9.1 must be replaced with other suitable forms. Moreover, by following the experiment over a long time, it can be established whether one species outcompetes the other and forms a stable monoculture or whether a stable coexistence results. A drawback of the Dewit experimental design is that there is only one total density, despite the range of proportions. It is therefore useful to repeat the experiment with different initial total densities, akin to altering the different initial conditions which led to the different monocultures in strong competition.

Competition and invasion of exotic species

Species have been invading new habitats ever since life first appeared and began to spread. As soon as there were two species, conditions were ripe for a situation in which either one invaded a habitat already occupied by the other. More recently, species in northern and previously glaciated areas are still adjusting their ranges in response to climatic warming since deglaciation, and much of palynology is a documentation of the timing and nature of changes in plant communities as species expand their ranges. Since range expansions were neither temporally or spatially synchronous, invasions of communities by “new” species are common in the pollen record. Most recently, anthropogenically induced global warming is raising the prospect of massive expansions and contractions of species ranges, with species invading communities and habitats where they are now absent. In addition, the massive economic globalization that is now occurring is inadvertently providing various species new transportation routes as “stowaways” in cargo. If the “exotic” species can invade a community of natives, then large changes in community and ecosystem structure and function can result: witness the invasion of eastern North American forests, wetlands, and lakes by gypsy moth, purple loosestrife, and zebra mussel, respectively, in the most recent years. Once these species successfully invade, they cause large and even permanent changes in the state of the community: in essence, the community undergoes a regime shift because it crosses a bifurcation to a system in which the invader can persist.

It is interesting to ask whether the Lotka–Volterra model can help us understand under what conditions a species can invade a habitat occupied by another. Yodzis (1989) reasons as follows: Suppose that a community is composed of species 2 near its carrying capacity, or $N_2 = K_2$, and species 1 is rare, or $N_1 = \varepsilon$. Under what conditions can N_1 increase? This can only happen when the growth rate for $N_1 = \varepsilon$ is positive when $N_2 = K_2$, or:

$$\left. \frac{dN_1}{dt} \right|_{N_1=\varepsilon} = r_1 \varepsilon (K_1 - \varepsilon - \alpha_{12} K_2) > 0$$

This can only happen when the term in parentheses is positive, or when $K_1 > \alpha_{12} K_2$. This happens when N_1 is a strong competitor against N_2 (and therefore its nullcline lies entirely above that of N_2), or when N_1 is a weak competitor, but not so weak as to have its nullcline lie entirely below that of N_2 . In the latter case, the two species will coexist, requiring $K_2 > \alpha_{21} K_1$. Mutual invasibility therefore usually implies mutual coexistence (Yodzis 1989, but see Armstrong and McGehee 1980 for some exceptions).

We often think of exotic invaders as “outcompeting” native species because of their rapid population growth during the initial stages of invasion, but invasion

can also be successful if the invader competes just strongly enough to maintain a positive population growth towards a coexistence equilibrium. And, of course, invasion can be initiated but be unsuccessful if the nullcline of the native species always lies above that of the invader, or if the initial invading population is too small to force the system across the separatrix where both species are strong competitors. The problem with unsuccessful invasions is that they leave little record behind: it is difficult to show that something has not happened without evidence or a long-term record of occupancy followed by apparent disappearance. Many of the current studies of exotic invaders have not yet been able to distinguish amongst these different modes of invasion, partly because it is not clear who exactly are the competitors of the invader and for what resource. The obvious experiment would be to release some exotic invaders under controlled conditions but of course such experiments would be difficult to justify: what if the invasion was successful? Here is a situation where we may have to rely on careful parameterization of appropriate competition models to make our best forecast of the probable success of invasion by an exotic species.

Mutualism

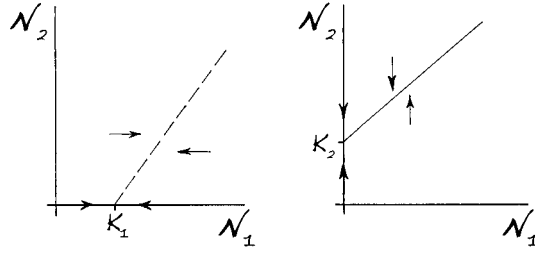
There are numerous examples of pairs of species that form mutualistic associations, where one or the other species or both benefits from the other. Such benefit may happen through increased access to soil nutrients by mycorrhizae when subsidized with carbohydrates from the host plant, or the eating of ticks on rhinoceroses by the oxpecker, and many other examples (Boucher 1985). But unlike competition, which involves a negative feedback between two species, mutualism involves a positive feedback which is potentially destabilizing. If I benefit you and you benefit me, then it is possible for our mutual benefits to increase our growth at ever increasing rates.

There are two broad types of mutualism, facultative and obligate. Facultative mutualists can live independently of each other, but their growth is enhanced in the presence of each other. Seed dispersal mutualists, such as the Clark's nutcracker and whitebark pine, are examples of facultative mutualists. Obligate mutualists cannot survive except in association with each other. Many digestive mutualists, such as the bacteria in the digestive tracts of ruminants and termites that actually digest the cellulose in the animal's food, are examples of obligate mutualists. Obligate mutualists are often so tightly linked that we think of them as one species. When we see a moose, we think of it as one organism, not as a symbiosis between moose and bacteria. Similarly, many lichens are obligate mutualisms between a fungus and an algae, and are classified as if the lichen is a species onto itself (the species name given to the lichen symbiosis is by convention the species name of the fungus rather than the name given to the algae associated with it).

The Lotka-Volterra competition model can be modified to serve as a beginning model for mutualism (Vandermeer and Boucher 1978) by making the interaction coefficients (α_{ij}) enhance growth rates rather than subtract from them:

$$\begin{cases} \frac{dN_1}{dt} = r_1 N_1 \left(\frac{K_1 - N_1 + \alpha_{12} N_2}{K_1} \right) \\ \frac{dN_2}{dt} = r_2 N_2 \left(\frac{K_2 - N_2 + \alpha_{21} N_1}{K_2} \right) \end{cases} \quad (9.10)$$

Fig. 9-5 Nullclines for N_1 (left and dashed line) and N_2 (right and solid line) for the Lotka-Volterra mutualism model, Eq. 9.10.



The nullclines for N_1 are:

$$N_1 = 0$$

$$N_1 = K_1 + \alpha_{12}N_2$$

Similarly, the nullclines for N_2 are:

$$N_2 = 0$$

$$N_2 = K_2 + \alpha_{21}N_1$$

In contrast to the competition model (Eq. 9.1), the nullclines are now increased from K_i in proportion to the product of N_j and α_{ij} . The mutualism coefficients therefore slant the nullclines in the opposite direction from the way that the competition coefficients slanted the nullclines (Fig. 9-5). But now, if the mutualism coefficients are large, then the nullclines may be slanted so much that they diverge in the first quadrant and stable coexistence is not possible. Let's analyze Eq. 9.10 for both facultative and obligate species to understand what determines whether two species can form a stable mutualism.

First, let's calculate the equilibria, which are:

$$N_1^* = N_2^* = 0 \quad (9.11a)$$

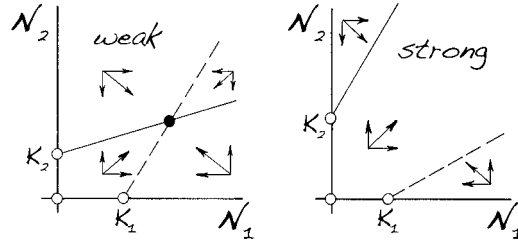
$$N_1^* = K_1, N_2^* = 0 \quad (9.11b)$$

$$N_1^* = 0, N_2^* = K_2 \quad (9.11c)$$

$$N_1^* = \frac{-K_1 - \alpha_{12}K_2}{\alpha_{12}\alpha_{21} - 1}, N_2^* = \frac{-K_2 - \alpha_{21}K_1}{\alpha_{12}\alpha_{21} - 1} \quad (9.11d)$$

Note the similarity with the equilibria of the competition model (Eq. 9.3), except for the sign changes in the coexistence equilibrium (Eq. 9.11d). In parallel with how we defined weak and strong competition, let's now define weak mutualism as that which happens when $\alpha_{12}\alpha_{21} < 1$ and strong mutualism when $\alpha_{12}\alpha_{21} > 1$. Let's also define facultative mutualism as having parameters $r_1 > 0$, $r_2 > 0$, $K_1 > 0$, $K_2 > 0$. That way, in the absence of the other species, the equilibrium population density of each species will be its carrying capacity, K_i . Similarly, let's define obligate mutualism as having parameters $r_1 < 0$, $r_2 < 0$, $K_1 < 0$, $K_2 < 0$. In the obligate case, the mutualism coefficients must be large enough to make the growth of each species positive in

Fig. 9-6 Nullclines, equilibria and their stability, and the vector fields for weak (left) and strong (right) facultative mutualism.



the presence of the other species, but each species on it's own will not have positive equilibria.

First, let's examine mutualism between two facultative species (Fig. 9-6), both having positive values for r_i and K_i (facultative-facultative mutualism). For coexistence to be possible, the equilibrium solution must lie in the first quadrant. By inspection of the coexistence equilibrium (Eq. 9.11d), this is only possible if $\alpha_{12}\alpha_{21} < 1$, which is weak facultative mutualism. Weak mutualism means that the positive feedback between the two species is not strong enough to result in strong, destabilizing positive feedback. On the other hand, if $\alpha_{12}\alpha_{21} > 1$ so that we have strong mutualism, then the coexistence equilibrium is in the third quadrant (both N_1^* and N_2^* are negative). This is because the nullclines in the first quadrant diverge from each other, which then causes them to intersect in the third quadrant. So, for two facultative mutualists to coexist, mutualism must be weak (Fig. 9-6).

Can two obligate mutualists coexist? By inspection of the coexistence equilibrium (Eq. 9.11d) for $K_1 < 0$ and $K_2 < 0$, a solution in the first quadrant is only possible if $\alpha_{12}\alpha_{21} > 1$, or if obligate mutualism is strong enough to overcome the negative intrinsic growth rates and carrying capacities. Weak obligate mutualism has an equilibrium solution in the third quadrant because its nullclines diverge in the first quadrant.

But are these coexistence equilibria stable? The vector field for the weak, facultative, mutualism coexistence indicates that the coexistence equilibrium is stable, whereas the vector field for the strong, obligate, mutualism coexistence indicates that the coexistence equilibrium is a semistable saddle-point. Let's see if we can verify this by examining the stability of the Jacobian evaluated at each of the coexistence points. The Jacobian is:

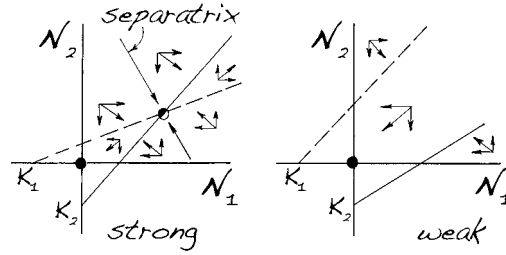
$$J|_{\text{coexistence}} = \begin{bmatrix} \frac{r_1(K_1 + \alpha_{12}K_2)}{K_1(\alpha_{12}\alpha_{21} - 1)} & -\frac{r_1\alpha_{12}(K_1 + \alpha_{12}K_2)}{K_1(\alpha_{12}\alpha_{21} - 1)} \\ -\frac{r_2\alpha_{21}(K_2 + \alpha_{21}K_1)}{K_2(\alpha_{12}\alpha_{21} - 1)} & \frac{r_2(K_2 + \alpha_{21}K_1)}{K_2(\alpha_{12}\alpha_{21} - 1)} \end{bmatrix} \quad (9.12)$$

Compare Eq. 9.12 with the Jacobian for the coexistence equilibrium in the competition model (Eq. 9.9).

The trace and determinant of the Jacobian at the coexistence equilibrium is:

$$\text{trace}(J|_{\text{coexistence}}) = \frac{(r_1 + r_2)K_1K_2 + r_1\alpha_{12}K_2^2 + r_2\alpha_{21}K_1^2}{K_1K_2(\alpha_{12}\alpha_{21} - 1)}$$

Fig. 9-7 Nullclines, equilibria and their stability, and the vector fields for strong (left) and weak (right) obligative mutualism.



$$\det(J|_{\text{coexistence}}) = \frac{-r_1 r_2 (K_1 + \alpha_{12} K_2)(K_2 + \alpha_{21} K_1)}{K_1 K_2 (\alpha_{12} \alpha_{21} - 1)}$$

For stability, the trace must be negative and the determinant positive. By inspection, for the weak facultative–facultative coexistence in the first quadrant (when $\alpha_{12}\alpha_{21} < 1$), the trace is negative and the determinant is positive and the coexistence equilibrium is a stable node. On the other hand, for the strong obligate–obligate coexistence in the first quadrant (when $\alpha_{12}\alpha_{21} > 1$), both the trace and the determinant are negative, and so the coexistence equilibrium is a semistable saddle-node.

Inspection of the vector fields (Figs 9-6 and 9-7) suggests that, as for the two-species competition models, limit cycles are not possible in facultative–facultative or obligate–obligate mutualism models: once the trajectory enters the regions between the two nullclines, it is either attracted to an equilibrium or it diverges to infinity. Again, the structure of the vector field appears to prohibit the trajectory from leaving these regions once it enters them, and so limit cycles do not appear possible. This can be proved algebraically from the signs of the elements of the Jacobian matrix, as we did above for competition models.

What about facultative–obligate associations? Let's designate N_1 as the facultative species ($r_1 > 0$, $K_1 > 0$) and N_2 as the obligate species ($r_2 < 0$, $K_2 < 0$). By applying these signs, the coexistence equilibrium becomes:

$$N_1^* = \frac{\alpha_{12} K_2 - K_1}{\alpha_{12} \alpha_{21} - 1}, \quad N_2^* = \frac{K_2 - \alpha_{21} K_1}{\alpha_{12} \alpha_{21} - 1}$$

For strong mutualism ($\alpha_{12}\alpha_{21} > 1$), the facultative–obligate coexistence is in the first quadrant when $K_1 < \alpha_{12} K_2$ and $K_1 < K_2/\alpha_{21}$. For weak mutualism ($\alpha_{12}\alpha_{21} < 1$), the facultative–obligate coexistence is in the first quadrant when $K_1 > \alpha_{12} K_2$ and $K_1 > K_2/\alpha_{21}$. If N_1 is facultative and N_2 is obligate, it can be seen by inspection that if $\alpha_{12}\alpha_{21} < 1$ and the mutualism is weak, the determinant of the Jacobian is positive and its trace is negative only when $K_1 > \alpha_{12} K_2$ and $K_1 < K_2/\alpha_{21}$, which puts the coexistence equilibrium in the first quadrant. Weak facultative–obligate mutualism, therefore, has a stable coexistence equilibrium in the first quadrant. By similar arguments, stability of coexistence in the first quadrant is not possible for strong facultative–obligate coexistence.

The conditions for stability of Lotka–Volterra mutualism models are stringent because the benefits of mutualism are linear – that is, growth increases in proportion to the population density of the cooperating species without a maximum. Various authors have tried different ways to limit the positive feedback of mutualism by having its

benefits operate according to some sort of diminishing returns. Wolin and Lawlor (1984) derive a model of mutualism by going back to the density-dependent assumptions of the logistic equation, and assume that the negative density-dependent term in the per capita birth rate equation for each species is inversely proportional to the population density of the mutualist. This bends the isoclines so that they do not diverge in the first quadrant: even under strong facultative mutualism there is then a stable coexistence (see also Kot 2001). Graves et al. (2006) also offer a similar approach, but based on having mutualism work by directly modifying r_i . The benefits of mutualism can also be constrained by explicitly including a resource whose availability to each species is enhanced by the mutualism, but whose availability is constrained by input–output mass balance (Lee et al. 1976). In later chapters, we will consider further how mass balance of resources constrains species interactions.

Exercise 9.1

Sketch the vector fields for all combinations of strong and weak facultative–obligate models. As with the competition, facultative–facultative, and obligate–obligate models, do the structures of these also suggest that limit cycles are not possible?

Exercise 9.2

In Chapter 8, we saw that a predator–prey system can be stable if only one species, the prey, has a finite carrying capacity. Can competitors and mutualists coexist stable even if only one has a finite carrying capacity? To find out, analyze the equilibria and their stability for the following competition model:

$$\begin{cases} \frac{dN_1}{dt} = r_1 N_1 \left(\frac{K_1 - N_1 - \alpha_{12} N_2}{K_1} \right) \\ \frac{dN_2}{dt} = r_2 N_2 - \alpha_{21} N_1 N_2 \end{cases}$$

and for the following mutualism model for both facultative–facultative and obligate–obligate mutualism.

$$\begin{cases} \frac{dN_1}{dt} = r_1 N_1 \left(\frac{K_1 - N_1 + \alpha_{12} N_2}{K_1} \right) \\ \frac{dN_2}{dt} = r_2 N_2 + \alpha_{21} N_1 N_2 \end{cases}$$

Do both species need to have a finite carrying capacity in the absence of each other for these models to have a stable coexistence equilibrium? Or does a finite carrying capacity for one suffice?

Summary: what have we learned?

In two-species competition models, the species reduce the growth of the other in proportion to their densities. Strong competition (values of the competition coefficients greater than some critical value) results in monocultures of one or either species. Stable coexistence is possible only with weak competition (values of the competition coefficients less than some critical value). Moreover, the conditions for monoculture and coexistence determine whether an invasion of one monoculture by the other is feasible. One striking feature of this model is the high degree of symmetry of the

differential equations and the solutions. In fact, other competition models which also possess this same symmetry also have the same conditions for stable coexistence. Density-dependent self-regulation, weak competition, and symmetry seem to be general properties of competition models that lead to similar qualitative behaviors.

Similarly, for Lotka–Volterra mutualism models, stable coexistence is possible only if at least one of the species is facultative (can survive without the other) and if the interaction coefficients are also weak. In the previous chapter on predator–prey interactions, we saw that stable coexistence is also possible only when the interaction strength between the predator and the prey is weak. So, too, for competition and mutualism: stable coexistence in Lotka–Volterra models is only possible when the interactions are weak. Therefore, in general, in two-species interactions models with Lotka–Volterra type coefficients, coexistence is possible only if the interactions between species are weak. Strong interactions between species appear to destabilize food webs.

Finally, the structure of the vector field and the signs of the elements of the Jacobian matrices prohibit complex eigenvalues for competition and mutualism models. Stable limit cycles are therefore not possible in two-species competition and mutualism models, as they are with two-species predator–prey models. In three-species competition models, however, stable limit cycles are possible (Gilpin 1975). In Chapter 11 we shall see that stable limit cycles between two competitors are possible if we explicitly include the resource for which they are competing and if certain other conditions on the competition terms are met. This puts the model into three dimensions (one for the resource and two for the competing species).

Open questions and loose ends

The requirement that both competition and mutualism be weak for stable coexistence in Lotka–Volterra models and also for various variations of the competition models (Ayala et al. 1973) raises some interesting but as yet unanswered experimental and philosophical issues: namely, how are we to distinguish weak interactions from no interactions? The competition and mutualism coefficients are the slopes of the nullclines and can be measured as the slopes of lines fitted to data (e.g., using a DeWit replacement experiment, which can also be done for facultative mutualism). If competition and mutualism are weak (as the theory requires for stable coexistence) then the slopes will not be very different from 0. But a slope of 0 is the expectation of the null hypothesis of no competition. Given the amount of noise in most ecological data, it may be difficult to detect a slope close to but statistically different from 0 unless we strongly control environmental conditions (such as climate, soil type, water chemistry, topography, etc.). Doing this requires very simplified conditions, such as common garden experiments (e.g., Wedin and Tilman 1993) or precisely controlled laboratory conditions (e.g., Ayala 1972), to minimize the noise in the data. But then the experimental conditions probably do not entirely match what is happening “in the field.” The generality of the experimental results is then somewhat restricted. On the other hand, if one wishes to do the experiment “in the field” with noisy data, then the common practice of using low probability levels (e.g., $P < 0.05$) to minimize Type I errors of rejecting the null hypothesis when it is true causes one to run the risk of making a Type II error, or accepting the null hypothesis when it is false. We therefore run the risk of rejecting the hypothesis of weak competition or mutualism as predicted by the theory when in fact it is true. Experimental design in ecology has always favored minimizing Type I errors, but it seems that detecting

competition or mutualism also requires us to consider the possibility and consequences of making Type II errors. Is the standard $P < 0.05$ for rejecting the null hypothesis too restrictive for detecting real but necessarily weak competition and mutualism coefficients? How should we design competition experiments to minimize making Type II errors so that we don't reject the conclusion that weak competition results in stable coexistence when it is true? Perhaps the long history of debate in ecology over whether competition happens and whether it is important in community dynamics has its origin partly in this experimental issue.

Throughout this chapter, we have been begging the questions: what resource are the species competing for or enhancing access to? Does it matter if the forms of the competitive or mutualism interactions are not symmetrical? How do the dynamics of the resource itself determine the behavior of the two species? We will explore these questions in a later chapter, but let's say at the outset that much work on these questions remains open.

What if one of the competitors or mutualists "engineers," or manages to create, its own environment (Jones et al. 1994)? Gurney and Lawton (1996) present a differential equation model of an ecosystem engineer and its habitat (resource) modifications that includes stable equilibria as well as cycles. It would be interesting to see how adding a competitor or a mutualist to this model affects its stability and the conditions for coexistence.

Finally, human societies harvest not only predator-prey systems, we also harvest systems of two or more competing or mutualistic species. The outcome of harvesting two competitors often leads to the less harvested one gaining a competitive edge over the other. Brauer and Castillo-Chavez (2001) present a good introduction to this problem, but a complete analysis of how the different harvesting functions presented in Chapter 7 affect the stability of two-species competition or mutualism systems remains to be done. What happens if the harvesting functions applied to each species are of the same form, thus preserving the symmetry of these equations, and what happens if they are of different forms, thus breaking the symmetry? What management implications result from these considerations?

10 Multispecies community and food web models

Generalized Lotka–Volterra models of communities

All communities contain more than two species. The extension of two-species models to n -species community or food web models comprises a vast literature, for which Pimm (2002) is the standard reference. Although in this chapter we cannot hope to cover the entire ground of these models, there are some points that must be made before we consider interactions of species with inorganic nutrients in the following chapters. We will also learn some new techniques for analyzing stability of n -species models when analytical solutions for eigenvalues cannot, in general, be explicitly found for $n \geq 5$.

Let's construct a general model for a community of n competitors. Recall that the competition equation for a particular species in a two-species Lotka–Volterra model looks like this:

$$\frac{dN_1}{dt} = N_1 \left(r_1 - \frac{r_1 \alpha_{11}}{K_1} N_1 - \frac{r_1 \alpha_{12}}{K_1} N_2 \right)$$

(in Eq. 9.1, we implicitly assumed that $\alpha_{11} = 1$, but here we will make it explicit so that in a moment we can write a general form for all species). Note that the coefficient of each competition term (including competition between individuals of the same species) is of the form $-r_i \alpha_{ij}/K_i$. We can then generalize this equation to include competitive effects from n species by adding terms of the form $(-r_i \alpha_{ij}/K_i)N_j$ for $j = 1 \dots n$ for each of n species of the same form:

$$\frac{dN_i}{dt} = N_i \left(r_i + \sum_{j=1}^n \frac{-r_i \alpha_{ij}}{K_i} N_j \right) \quad (10.1)$$

We can write this system of equations more compactly in matrix-vector notation as:

$$\frac{d\mathbf{N}}{dt} = \mathbf{N}(\mathbf{p} + \mathbf{A}\mathbf{N}) \quad (10.2)$$

where $\mathbf{N}$ is a vector of the population densities of each of n species, $\mathbf{p}$ is a vector of each r_i , and $\mathbf{A}$ is a matrix whose entries are $-r_i \alpha_{ij}/K_i$. This is now a generalized Lotka–Volterra competition model.

The matrix $\mathbf{A}$ is sometimes called the community matrix. If the interaction terms have positive sign, then they represent a community of mutualists. If the interaction

terms between two species in the generalized Lotka–Volterra model are not symmetrical, as they are in competitive or mutualistic interactions, but instead represent predator–prey interactions, then we have a food web model of who eats whom. Therefore, by substituting other appropriate coefficients in place of r_i and $-r_i\alpha_{ij}/K_i$, we can simulate any combination of competition, mutualism, and predation for n species (if a pair of species does not interact, then the coefficient is zero). We now have a general model for the dynamics of an n -species community in which any pair of species can interact via predation, competition, or mutualism, if they interact at all.

How does the community matrix, A , relate to the Jacobian matrix, J ? Write Eq. 10.1 more generally as:

$$f_i = \frac{dN_i}{dt} = N_i \left(r_i + \sum_{j=1}^n a_{ij} N_j \right)$$

where a_{ij} are the forms of the coefficients appropriate for each two-species interaction and the elements of A . Recall that the elements of the Jacobian are the partial derivatives of f_i with respect to each species. Therefore, for $i \neq j$, the off-diagonal elements of the Jacobian are:

$$\frac{df_i}{dN_j} = a_{ij} N_i$$

And the partial derivatives of each species with respect to itself (the diagonal elements of the Jacobian) are:

$$\frac{df_i}{dN_i} = r_i + 2a_{ii}N_i + \sum_{j=1, j \neq i}^n a_{ij}N_j$$

Evaluation of the Jacobian at an equilibrium can be accomplished simply by inserting the equilibrium solutions for N_i into the above equations. More compactly, the Jacobian can be calculated as the community matrix times a diagonal matrix where the diagonal elements are the equilibrium solutions for N_i . Because the Jacobian is a linear transformation of the community matrix, whatever conclusion about stability that applies to the Jacobian also applies to the community matrix. So two-species competitive interactions can be extended to n species by simple generalization of the mathematics of Jacobian matrices. Except for the signs of the interaction terms, similar models can be constructed for multispecies mutualism communities.

The stability of community and food web models

Community and food web stability and the conditions that determine it have been and remain central questions in ecology (Pimm 2002). Can we analyze the stability of n -species community models in the same way as for two-species models? How does community stability change with increased number of species? To some extent, this depends on how many species interact with each other, on the number of links between all pairs of species, and the structure of the interaction network. Are more species-rich community models more likely to be stable? Does increasing the number of links between species increase the chance of being stable?

In principle, we should be able to answer these questions by analytically solving the eigenvalues of the Jacobian of generalized Lotka–Volterra model of any arbitrary size, no? No. The eigenvalues become extremely complicated for $n = 4$ and they are impossible to solve analytically in general form for $n \geq 5$. The reason for this is that to solve for the eigenvalues, we must find the roots of an n -degree polynomial (recall that to find the eigenvalues of two-species models, we used the quadratic equation for the roots of a second degree polynomial). But in 1824, Niels Henrik Abel (1802–1829) showed that it is not possible, in general, to find analytical solutions for the roots of polynomials of degree 5 or higher (see Pesic 2003 for a history of the problem, an English translation of Abel’s paper, and a brief biography of Abel, who lived a romantically tragic life). For some special cases with rather severe restrictions on the forms of the coefficients, it is possible, albeit difficult, to find analytical solutions for eigenvalues, but they are almost always too complicated to interpret.

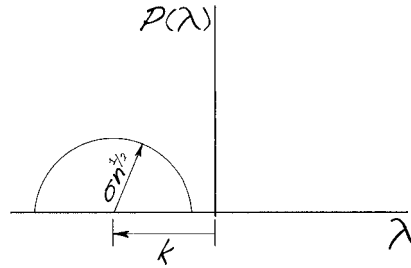
But all is not lost. We can approach the problem of the stability of community models in several ways. First, we can simulate the trajectories of the system of equations for particular parameter values, using Runge-Kutta numerical methods outlined in Chapter 2. Gilpin (1979), Schaffer et al. (1986), and Hastings and Powell (1991) all show, mainly by constructing numerical solutions of trajectories, that three-species models apparently possess strange attractors and therefore chaotic behavior.

Second, even though Abel’s results prevent us from finding general expressions for eigenvalues of symbolic matrices of size $n \geq 5$, if we assign numerical values to the entries of a Jacobian matrix for any particular community, we can then find numerical values of its eigenvalues using various software programs for numerical solutions (see Appendix for some Matlab commands to do this). This will tell us whether that particular community model is stable or not. We can then construct a large number of community models for various parameter values and see what proportion of these matrices have all of their eigenvalues with negative real parts and hence are stable.

Third, we can derive some qualitative rules for determining the likelihood of food web stability given only that each species positively or negatively affects the other species or has no effect.

To start analyzing these models along these lines, we ask what should be the null hypothesis of a Jacobian matrix of a community? One possibility is that each element is drawn from a random normal distribution with a mean of 0 and standard deviation σ . The standard deviation of the elements is a measure of the strength of the interactions between any two species: the higher the standard deviation, the greater the probability that the interaction between two species is strong (very positive or very negative). A small standard deviation means that the elements are clustered around zero, and hence weak. Let’s put some additional conditions on the structure of the Jacobian, and then proceed by relaxing these assumptions. Symmetric, random matrices are nice to start with because all their eigenvalues are real (see Searle 1966, p. 192, for proof) and their eigenvalues have some simple statistical distributions. The elements of a symmetric matrix are symmetric about their diagonals, so $a_{ij} = a_{ji}$. This means that for any pair of species, their competitive or mutualistic effect on each other are numerically equal. Suppose then we have a community of n species whose structure is described by a random symmetric matrix with entries drawn from a normal random distribution with mean 0 and standard deviation σ . If the number of species is very large, then the eigenvalues will approximate a continuous

Fig. 10-1 The frequency distribution of eigenvalues of a random, real symmetric matrix with all diagonal elements = k (see Wigner 1959 and Post et al. 1978).



distribution. Wigner (1959) shows that the frequency distribution of real parts of eigenvalues of such matrices is a semicircle with radius $\sigma n^{1/2}$ (see also Mehta 1967). Half of the eigenvalues will have positive real parts and half will have negative real parts. But if we now subtract a constant k from each diagonal entry, the semicircle will be shifted to the left by that amount and the probability of obtaining eigenvalues with negative real parts will increase (Fig. 10-1). So, if $\sigma n^{1/2} < k$, then all eigenvalues will usually have negative real parts and the food web will likely be stable (May 1972b, Post et al. 1978). This result is also true if the elements are drawn from a random uniform distribution instead of from a random normal distribution.

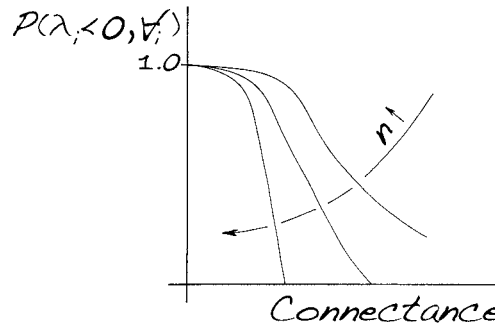
What is the biological meaning of having all the diagonal elements of the Jacobian decreased by k ? It is this: k measures the negative effect of each species populations on itself through density dependent self-regulation. But we also know that density-dependent self-regulation implies that each species has a carrying capacity K . So, the interaction strengths between each of n species must be weaker than the negative interactions between individuals within each species that underlie density-dependent self-regulation. Therefore, weak interspecies interactions increase the probability of a random community model being stable, and the more species there are in a community model, the weaker the interactions have to be to increase the probability of being stable. (This is also true of n species competition and mutualism community models as well (McCann et al. 1998).) We have seen that the stabilities of two-species predator-prey, competition, and mutualism models depend on various inequalities between K and the parameters that control interactions between the species, and that the stability of co-existence in all three models is conferred by weak interactions relative to density dependent self-regulation responsible for carrying capacities. This appears to hold in general for larger community models as well (Yodzis 1981).

A caveat: not all of the diagonal elements of the Jacobian need to be negative. Recall the Jacobian for the predator-prey model with logistic growth for the prey in the absence of the predator (Eq. 8.13):

$$J|_{\text{coexistence}} = \begin{bmatrix} -\frac{rm}{\beta h K} & -\frac{m}{\beta} \\ \beta r \left(1 - \frac{m}{\beta h K}\right) & 0 \end{bmatrix}$$

Here, the effect of the predator on itself is zero, but the coexistence equilibrium is stable because the trace is still negative. But on the whole, the more diagonal elements that are negative, the more likely the community model is to be stable (Yodzis 1981).

Fig. 10-2 Increased species richness (n) and increased connectance reduces the probability of stability of a community model.



How does the complexity of the community model affect the probability of its stability? We can define complexity as the proportion of all possible interactions between species that are realized. This is known as the system's connectance, C . The total possible number of interactions between n species is n^2 . If L is the number of realized links in a community, $C = L/n(n - 1)$. Gardner and Ashby (1970) simulated a large number of matrices with different connectances and different numbers of species and showed that the probability of stability decreases rapidly once some value of connectance is exceeded; furthermore, both this critical value and the probability of stability decrease more rapidly as the number of species increases (Fig. 10-2).

May (1972b) relaxed the assumption of symmetric matrices by allowing some random proportion of the Jacobian's elements to be 0, and McMurtrie (1975) also considered asymmetric matrices where $a_{ij} = -a_{ji}$. In both studies, the diagonal elements had value $k = 1$. Both concluded that community models are usually stable if:

$$\sigma(Cn)^{1/2} < k$$

Since $C = L/n(n - 1)$, this condition can be rewritten as:

$$\sigma\left(\frac{L}{n-1}\right)^{1/2} < k$$

It can be argued that communities in nature are not random assemblages of species interactions: indeed, natural selection selects from random mutations and produces nonrandom interactions. But Pimm (1979) and Yodzis (1980) analyzed biologically realistic community matrices and found qualitatively similar results as shown in Fig. 10-2.

So, in general, increased number of species must be accompanied by decreased strength of species interactions for community models to be stable, and the probability of stability is enhanced by each species having strong density-dependent self-regulation, implying some sort of carrying capacities for each species.

Community models that are entirely donor-controlled form notable exceptions to these conclusions. Donor control means that transfer of matter and energy between species is entirely under the control of the donating species and not controlled by the recipient species. The transfer might be modeled by a linear transfer of a

proportion of the material from the consumer to the predator, for example. DeAngelis (1975) shows that for donor-controlled models, increased connectance and species number increases the probability of the model being stable, in contrast to the above results in Fig. 10-2. But completely donor-controlled models are often not very realistic models of communities or ecosystems because the transfer of materials and energy usually depends on the behavior and population densities of both the recipient as well as the donor species. For example, except for the quota harvesting model, all of the functions derived in Chapter 7 for uptake, consumption, or harvesting of one resource or species by another species depend on population densities of both interacting species.

What if we do not know the strength of interactions between species, but only their sign? That is, what if all we know is whether two species form predator–prey interactions in which the prey has a positive effect on the predator and the predator has a negative effect on the prey, competitive interactions in which each species has a negative effect on each other, mutualistic interactions in which each species has a positive effect on each other, or if they do not interact? For most communities, this is all we know, based largely on knowledge of natural history of each species and perhaps analysis of stomach contents. Can we draw any conclusions of the stability of community models based only on qualitative information on the signs of species interactions?

It turns out we can, based on a theorem from economics stated by Quirk and Ruppert (1965; see also May 1975b). They state the following necessary and sufficient conditions for a matrix to be stable, based only on the signs of interactions:

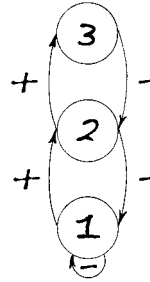
- (i) $a_{ii} \leq 0$ for all i .
- (ii) $a_{ii} \neq 0$ for at least one i .
- (iii) $a_{ij}a_{ji} \leq 0$ for all $i \neq j$.
- (iv) For any loop of three or more species $i, j, k, \dots, q, r$, the product of the successive interaction coefficients $a_{ij}a_{jk} \dots a_{qr}a_{ri} = 0$, $i \neq j \neq k \dots \neq q \neq r$.
- (v) $\det J \neq 0$.

If a matrix meets these conditions, then it is called “qualitatively stable.” If it does not, whether or not it is Lyapunov or asymptotically stable depends on the exact forms of the elements and the parameter values. Therefore, if we can determine if a matrix is qualitatively stable by examining only the signs of the elements, then it is stable in the sense that all of its eigenvalues have negative real parts (asymptotic Lyapunov stability). But if it does not pass the Quirk–Ruppert criteria, then we have to calculate the eigenvalues and examine them for bifurcations with different parameter values.

What are the biological meanings of these conditions? Condition (i) and (ii) mean that none of the diagonal elements are positive with at least one being negative, which as we have seen increases the probability of eigenvalues of random matrices having negative real parts. Therefore, at least one species has density-dependent self-regulation.

Condition (iii) means that predator–prey (+ −), commensal (0 +), and amensal (0 −) interactions are more likely to provide a model with stability than competitive (− −) and mutualistic (+ +) interactions because these latter two interactions give $a_{ij}a_{ji} > 0$. Of the three interactions that are most likely to confer stability, predator–prey interactions are by far the most common. Therefore, condition (iii) almost guarantees that stable food webs will have distinct trophic levels in which species

Fig. 10-3 A three-species, predator–prey food web.



in higher levels consume species at lower trophic levels (all of the remaining interactions can take place at the same trophic level).

Condition (iv) guarantees that there are no feedbacks that begin with any species, flow through a loop of three or more species, only to end by increasing the growth of the beginning species (thus causing that species population density to increase without bounds) or decrease the growth of the beginning species (thus driving that species to extinction). If the loop results in increasing the growth of the first species, then it is clearly destabilizing. If the loop ends in decreasing the growth of the first species, then it may be stabilizing unless it is strong enough to drive the first species to extinction. The detailed form of the interactions then needs to be analyzed to distinguish between the two possibilities. Condition (v) is a technical mathematical condition that prevents the matrix from being underdetermined – that is, that there will be as many equations as there are species.

Let's apply the Quirk–Ruppert conditions to some possible food web configurations. First, let's examine the qualitative stability of a three-trophic-level predator–prey model, with only the lowest trophic level having density-dependent self-regulation (Fig. 10-3). The qualitative interaction matrix is:

$$\mathbf{A} = \begin{bmatrix} - & - & 0 \\ + & 0 & - \\ 0 & + & 0 \end{bmatrix}$$

For this matrix:

- (i) All $a_{ii} \leq 0$.
- (ii) $a_{11} < 0$.
- (iii) All $a_{ij}a_{ji} \leq 0$ for all $i \neq j$.
- (iv) For three species, there are two loops. One loop is:

$$a_{12}a_{23}a_{31} = 0.$$

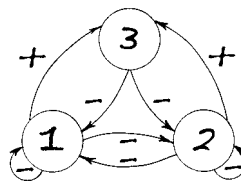
All other loops in this direction are permutations of this one. To find the set of loops in the other direction, simply transpose the i and j :

$$a_{21}a_{13}a_{32} = 0.$$

- (v) $\det \mathbf{A} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \neq 0$.

So three-trophic-level predator–prey models are qualitatively stable and are therefore stable in the sense of all eigenvalues having negative real parts. This is also usually true for n -species predator–prey food webs (Allesima and Pascual 2008).

Fig. 10-4 A three-species community, with two competitors and a predator.



What about models in which two species are competitors (which also have density-dependent self-regulation), with a consumer that eats both (Fig. 10-4)? This is a common case for many herbivores that have a mixed diet. The qualitative interaction matrix is:

$$A = \begin{bmatrix} - & - & - \\ - & - & - \\ + & + & 0 \end{bmatrix}$$

For this matrix:

- (i) All $a_{ii} \leq 0$.
- (ii) $a_{11} < 0$.
- (iii) $a_{12} a_{21} > 0$.
- (iv) The three-species loops are:

$$a_{12}a_{23}a_{31} = - + - = +$$

$$a_{21}a_{13}a_{32} = + - - = +$$

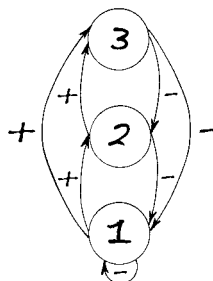
- (v) $\det J \neq 0$.

So conditions (iii) and (iv) fail because of the competition between species 1 and 2. This does not mean that the community is necessarily unstable, only that whether or not it is stable depends on the forms of these interactions and the parameter values and cannot be determined solely from the qualitative criteria. We need to know the specific equations for this community model and the eigenvalues of the Jacobian need to be evaluated to determine the conditions for stability.

Exercise 10.1

An omnivore is a species that feeds on more than one trophic level (Fig. 10-5). Bears are omnivores because they prey on moose as well as berries. Construct the qualitative interaction matrix from the following graph of a three-species system with the top predator being an omnivore and analyze whether it is qualitatively stable.

Fig. 10.5 A three-species food web in which species 3 is an omnivore.



Summary: what have we learned?

Two-species Lotka–Volterra models can be generalized to n -species community or food web models. But analytical solutions of eigenvalues cannot be solved in general for $n \geq 5$. Instead, two approaches can be taken for examining stability of large food webs. One is to simulate a large number of numerical realizations of particular food web matrices and calculate the eigenvalues numerically. The proportion of these that have all eigenvalues with negative real parts gives the probability that such a community or food web will be stable. Alternatively, one can determine qualitative stability using the Quirk–Ruppert criteria. If a particular qualitative matrix (entries +, –, or 0 to correspond to the sign of the interaction between pairs of species) passes these stability criteria, then it is stable in the sense that all eigenvalues have negative real parts (asymptotic Lyapunov stability). But if it does not pass these criteria, then the model may or may not be Lyapunov or asymptotically stable and we need to examine the equations and their Jacobian to determine stability.

Although there are exceptions in particular cases, studies of community and food web stabilities have converged on several general conclusions: (1) increasing the number of species will likely decrease stability of community or food web models; (2) increasing connectance of the community or food web also will likely decrease their stability; (3) decreased interaction strength between species increases the chance that community or food web models will remain stable with increased richness and connectance; (4) at least one species in the community or food web model needs to have density-dependent self-regulation for stability, and the more species that self-regulate the greater the chance of stability – this means that at least one species must possess a carrying capacity; (5) food web models composed of predator–prey interactions (which guarantee different trophic levels) are more likely to be stable than communities that also include competitive, mutualistic, or omnivorous interactions.

Open questions and loose ends

There are many open questions in food web structure and stability that cannot be covered here, but are covered very extensively in Pimm (2002). However, I would like to point out two problems of great interest to me, as examples.

The first problem is that of succession of communities or food webs. Succession implies instability of one food web as it is replaced by another; eventually a stable climax community may or may not be reached, or succession can be cyclic (Wein and El-Bayoumi 1983). Pimm (2002) gives a very interesting example of the problem that succession poses for food web models. Assume that two species are forming a predator–prey system which is invaded by a third that is omnivorous for the first two. Omnivory can destabilize food webs, and so one of the original species may go extinct after the invasion. Assume it is species 2. Can another omnivore now invade the species 1–species 3 food web? Are there particular food web structures that resist invasion by an “exotic” species? To what extent does this depend on how the exotic species interacts with the extant species – as a competitor, a mutualist, a prey, or a predator? Are there particular community structures or initial conditions that inevitably lead to cyclic succession, such as Wein and El-Bayoumi (1983) report for boreal forests, even without periodic disturbance?

Perhaps the most discussed question in food web theory these days is how does food web structure affect and depend upon the flux of materials and energy through

it? We have seen that increased species richness and connectance tends to decrease stability of food web models, but would these results be reversed if species interactions were constrained by mass balance of energy and nutrients? This is a very open question, although Thébault and Loreau (2003) have made a promising beginning. But this brings us to the next four chapters on interactions between species and inorganic resources.

Part 3

Ecosystems

11 Inorganic resources, mass balance, resource uptake, and resource use efficiency

Why consider inorganic resources?

So far, we have made remarkable progress in elaborating the structure and dynamics of ecological systems, beginning with the unstable exponential or geometric models of population growth and proceeding by identifying hidden assumptions in these models and relaxing them. By doing this and applying the techniques of eigenvalue stability and bifurcation analysis, we have been able to uncover a robust set of rules within which species can interact and either coexist stably or else drive others (or be driven) to extinction. In the end, we have been able to generate many of the rules of community and food web assembly for an arbitrary number of species. Using the techniques of bifurcation analysis, we have uncovered some critical values for parameters which generate sudden changes in behaviors, analogous to regime shifts in nature, once they are crossed.

In all of these models, there is always one parameter, the carrying capacity of the species that forms the base of the food chain or web, usually termed the primary producer, which appears in all the analyses for stability and the nature of the bifurcations. We saw this first in the various predator–prey models, but it reappeared in two-species competition and mutualism models as well as generalized Lotka–Volterra and food web models. Without at least one species having a carrying capacity, the models were unstable.

But K is simply the nontrivial equilibrium determined by the ratio of initial net population growth when population levels are low (r) to the sum of density-dependent rates of self-regulation of per capita birth and death ($\beta + \delta$). We have yet to explicitly identify and define the mechanism behind this self-regulation. As we pointed out in Chapter 5 and several times since, the carrying capacity, K , is a result of a particular assumption about density dependence and has nothing explicitly to do with the environment (see also Williams 1972). Williams (1972) also makes the point that because K makes no specific reference to any environmental factor it is difficult, if not impossible, to predict population behavior in different environments without first determining K , which requires observing the population in the particular environment in the first place! In our development of more and more complicated models of species interactions and food webs, we have simply pushed the problem of what precisely is the “environment” which is constraining species interactions down to the lowest species in the food web or community, the primary producers. The problem is still with us. What is the mechanism that constrains the growth of the lowest member of a food web or community and, by implication, every species that feeds upon it, competes with it, or cooperates with it?

The answer is that ultimately organisms are made of “stuff” and it is the mass balance of this stuff which constrains species dynamics (Williams 1972, Wedin 1994).

This stuff is a variety of abiotic (nonliving) materials – ultimately the elements, along with sunlight, that results in a chemical disequilibrium when it is captured by chlorophyll. The flow of energy in this open system therefore drives the transfer of stuff up the food chain. Not only that, organisms combine these abiotic materials only in certain ratios – indeed the ratios of different elements are almost an essential feature of each species (Sterner and Elser 2002). Once we are talking about elements and elemental ratios, we are entering the realms of physics and chemistry. In those fields, the conservation laws are paramount, and the conservation law that will concern us most is the conservation of mass. Conservation of mass constrains how the abiotic materials that are the stuff of living things flow through food webs. Mathematically, the constraints of conservation laws impose severe restrictions on the number and form of solutions to equations of ecological interactions (Pastor 2003).

In the next several chapters, we will explore how the “constraints of abiotic stuff” and its transfer within food webs and ecosystems affect stability and coexistence. We will see that mass balance of elemental resources through input–output budgets is the final solution to the problem of carrying capacity that has bedeviled us. In fact, mass balance of resources is ultimately what carrying capacity is all about. It is this idea that is the main contribution of ecosystem ecology to ecology in general.

Let’s begin first by exploring the behavior of a single abiotic resource, typically an inorganic nutrient but it can also be a vitamin (Droop 1968) or even an amino acid (Kielland 1994). We will begin by exploring the behavior of this resource in isolation from any interactions with primary producers or decomposers. We will then introduce the process of uptake of this resource and how uptake affects species interactions and coexistence.

Mass balance of a single inorganic resource

As soon as we begin to consider the mass balance of a resource, we must distinguish between open systems (with inputs and outputs) and closed systems with no outside flux to and from the surrounding environment. The dynamics of a closed resource pool of size R in isolation are trivial: R is constant and nothing happens. In later chapters when we add higher trophic levels we shall find that the consequences of closed systems are not so trivial.

The case of an open resource is more interesting (Fig. 11-1). Since the resource we will be considering is inorganic, it exerts no control over inputs, I , to it. Let’s assume that outputs or exports from the resource pool are lost at some fraction, l , of its density. The units of I are mass or volume per t and the units of l are mass or volume of output per mass or volume of R in the pool per time. Then the mass balance of inputs and outputs determines the rate of change of the density of the resource pool, just as the balance of births and deaths determines the rate of change of a living population:

$$\frac{dR}{dt} = I - lR \quad (11.1)$$

It should be readily apparent to you by now that the equilibrium resource density, R^* , equals I/l . It should also be obvious that the eigenvalue of the system must equal

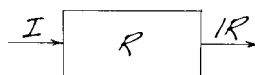


Fig. 11-1 The flow of a resource through a pool.

$-l$ at all values of R and that the system is therefore globally stable, which is simply a consequence of the fact that Eq. 11.1 is linear.

This equation is simple enough that we can solve it explicitly by separation of variables to obtain the time-dependent solution. This solution has some interesting implications. We begin by factoring out a $-l$ from the r.h.s., separate the parts in R and t , and integrate both sides:

$$\int \frac{dR}{R - \frac{I}{l}} = -l \int dt$$

which results in:

$$\ln\left(R - \frac{I}{l}\right) = -lt + C$$

The constant of integration, C , is evaluated as the initial condition when $t = 0$, or $C = \ln(R(0) - I/l)$. Inserting this into the above equation, we obtain:

$$\ln\left(R(t) - \frac{I}{l}\right) - \ln\left(R(0) - \frac{I}{l}\right) = \ln\left(\frac{R(t) - \frac{I}{l}}{R(0) - \frac{I}{l}}\right) = -lt$$

Now exponentiate both sides:

$$\frac{R(t) - \frac{I}{l}}{R(0) - \frac{I}{l}} = e^{-lt}$$

Now after a bit of algebra to isolate $R(t)$, which you should verify on a piece of scrap paper, we obtain:

$$R(t) = R(0)e^{-lt} + \frac{I}{l}(1 - e^{-lt}) \quad (11.2)$$

At $t = 0$, $R(t) = R(0)$. The first term on the r.h.s. therefore describes the fate of the initial amount of R , which decays exponentially as a constant fraction l of it is exported. However, more R enters through the input and a fraction of the resulting pool is also lost. The second term describes the fate of the new R entering the system. Therefore, as $t \rightarrow \infty$, $R(t) \rightarrow I/l = R^*$.

The terms in e^{-lt} give the transient approach of $R(t)$ towards the equilibrium. If we assume that the system is perturbed away from R^* to $R(0)$, then when $e^{-lt} = e^{-1}$ the perturbation or disturbance will have decayed to within a fraction $1/e$ of R^* . At this time t must equal the reciprocal of l , or $1/l$. This is, from Chapter 2, the characteristic return time of a system to equilibrium, or $1/|\lambda|$.

How long will a given atom or molecule of R reside in the pool when the pool is at equilibrium? This is given by the ratio of the equilibrium R^* to the output flux, lR^* , or again $1/l$. Therefore, the fractional loss rate, l , by virtue of also being the eigenvalue of the system, determines the stability of the equilibrium, the rate of return to equilibrium after a perturbation or disturbance, and the mean residence time of R in the system. The fractional loss rate is sometimes called the *resource turnover rate* or (assuming the resource in question is a nutrient) the *nutrient turnover rate*.

Exercise 11.1

A resource is really a general term for any inorganic constituent of an ecosystem. It need not be something beneficial. A pollutant also follows the same dynamics of a resource in that it has an input rate and a fractional output rate. Assume that at $t = 0$ the amount of a pollutant in a lake $P(0) = 0$. Now, assume that the pollutant begins entering the lake via the only inlet stream at a rate of 10 kg per week. The pollutant stays in the water column. The amount of pollutant in the water column at any time is P , in kg. Some of the water flows out the only outlet stream at a fractional rate 0.02 (liters/liters of total lake volume) $[-]$ wk^{-1} , and takes whatever pollutant is dissolved in it (we ignore evaporative losses). Since we are assuming that the pollutant is homogeneously distributed throughout the lake (why has this been an implicit assumption?), then the fractional export rate of the water is also the fractional export rate of the pollutant.

- 1 What is the amount of pollutant held by the lake when it reaches steady state?
- 2 How long would it take to reach one-quarter of that amount?
- 3 It is later discovered that, in addition to the pollutant leaving by the outflow stream, some of the pollutant also is absorbed permanently onto bottom sediment at the fractional rate of 0.005 wk^{-1} . Derive an equation for the amount of pollutant, P , at any time t with the same input as before but now with the two outputs (export in the stream and absorption onto the sediment) and recalculate parts (1) and (2) above given this new data.

Exercise 11.2

Burial is an important problem in global carbon cycling – organic matter on the seafloor is decaying, but at the same time it is also being buried under new sediment. What is buried is removed from interaction with the water column and eventually gets metamorphosed into fossil fuel. What gets buried is not a constant fraction of the organic matter but a constant thickness of sediment, or a constant volume (strictly speaking, this assumption holds if the concentration of organic matter in the water column is not negligible). Burial of a resource, in this case carbon, can be thought of as a negative input. If we assume that organic matter in the seafloor sediments is well mixed and we neglect additional inputs, then we can model the loss rate of a given amount of organic matter already in the seafloor sediments as:

$$\frac{dC}{dt} = -kC - B$$

- 1 Solve this equation by separation of variables as above for $C(t)$.
- 2 What is the limit of $C(t)$ as $t \rightarrow \infty$? Why is it a negative number?
- 3 Derive an expression for how long it takes for $C(t)$ to equal 0. This is the amount of time the organic matter will remain “unburied” and releasing carbon back into the water column by decomposition. The difference between $C(t) = 0$ and the limit from your answer to 11.2b is the amount of $C(0)$ that is buried. How does this vary with burial rate B and decay rate k ?

Resource uptake, carrying capacity, and the logistic model

We can now consider how uptake of a resource stabilizes biomass through mass balance. Williams (1972) provides two derivations, one for a closed system and another for an open system, both of which lead to the logistic model but now with mass balance explicitly considered. (Incidentally, this paper is a superb model of mathematical reasoning in ecology and while it may be difficult to find, it is well worth the effort.)

Williams begins by assuming that the uptake of a nutrient and subsequent growth of the primary producer or microbe (dB/dt) is proportional to the density of both the resource pool (R) and the biomass of the organism (B):

$$\frac{dB}{dt} = \beta u B R \quad (11.3)$$

where u is specific uptake rate of R per unit B and β is the net amount of B produced per unit R taken up, a measure of the efficiency by which R is converted to B (we will return to resource use efficiency at the end of the chapter). You should recognize this as the proportional harvesting function, or Lotka–Volterra predator–prey interaction between B and R . We also assume that the change in R depends at least partly on the conversion of R to B because of uptake:

$$\frac{dR}{dt} = -u B R \quad (11.4)$$

For a closed system, conservation of mass requires that the total amount of the resource contained in R and B be constant, or

$$R(0) + \beta B(0) = R + \beta B \quad (11.5)$$

where $R(0)$ and $B(0)$ are the initial pool sizes of R and B respectively. Because the system is closed, $R + B$ is constant and as B increases with uptake of nutrients, R must decrease in proportion, the efficiency term determining the proportionality. Solving Eq. 11.5 for R and substituting into Eq. 11.3, we obtain:

$$\frac{dB}{dt} = \beta u B \left[R(0) + \frac{B(0) - B}{\beta} \right]$$

And rearranging terms, we obtain:

$$\frac{dB}{dt} = \beta u \left[R(0) + \frac{B(0)}{\beta} \right] B - u B^2 \quad (11.6)$$

If this looks to you suspiciously like a logistic model, it is. Compare it with the following form of the logistic model:

$$\frac{dN}{dt} = rN - \frac{r}{K} N^2$$

With a bit of algebra, you can see the following formal equivalences between the parameters of Eq. 11.6 and the logistic model:

$$B \triangleq N, r \triangleq \beta u \left[R(0) + \frac{B(0)}{\beta} \right], K \triangleq B(0) + \beta R(0) \quad (11.7)$$

(the $\overset{\Delta}{=}$ symbol means “formal equivalence” of one term in one equation for a term in another equation that plays an identical mathematical role). So we have now defined r and K explicitly in terms of the initial conditions and the parameters that determined the transfer of R to B under the constraints of mass balance for a closed system. The formal equivalence with K implies that if we start with more B or R in the closed system, then eventually the resource will be distributed into a larger biomass at equilibrium. Similarly, if the efficiency of converting the resource into biomass increases, so will the equilibrium size of B . K is simply the result of the mass transfer of R to the plants to make B at an efficiency β .

What about for an open system in which there is input to R and output of both R and B ? This is actually a model for a chemostat. Assume that inputs, I are the product of an inflow rate of the medium carrying the resource (usually water), v , and the concentration of the resource in the inflow, C_i , both new parameters, not variables. Assume that the resource not taken up is exported in the same medium with the same influx rate (to maintain continuity of flow), but that its concentration has now changed to C because of uptake. (We assume that the organisms do not significantly affect the flux of the medium of transport, merely the resource within it. This would be the case for aquatic plants, such as algae.) Assume also that the plants die and are transported out of the system at a specific rate, d . The differential equations for C and B are now:

$$\begin{cases} \frac{dC}{dt} = vC_i - uBC - vC = v(C_i - C) - uBC \\ \frac{dB}{dt} = \beta uBC - mB \end{cases} \quad (11.8)$$

The change in concentration between C_i and C happens because of plant uptake and sequestering of the nutrient in biomass with an efficiency β . Mass balance then requires that

$$B = \beta(C_i - C) \quad (11.9)$$

As before, we solve this for C :

$$C = C_i - \frac{B}{\beta}$$

and substitute into the differential equation for B :

$$\frac{dB}{dt} = \beta uB \left[C_i - \frac{B}{\beta} \right] - mB$$

After a bit of algebra to expand the terms and gather all terms of the same power we obtain:

$$\frac{dB}{dt} = (\beta uC_i - m)B - uB^2 \quad (11.10)$$

As before, we can make the following formal equivalences with the logistic model:

$$B \triangleq N, \quad r \triangleq \beta u C_I - m, \quad K \triangleq \beta C_I - \frac{m}{u} \quad (11.11)$$

If we increase the concentration of the resource in the influx, C_I , the uptake rate, u , and the resource use efficiency, β , then K is increased but if the death/export rate of the biomass, m , increases, then K is decreased.

So we can recover a logistic model for open systems as well in which r and K are now explicitly defined in terms of more fundamental mass balance constraints of the environmental fluxes of the resource rather than an *ad hoc* assumption of linear dependence of per capita birth and death rates on population density without explicit regard to the environment which presumably caused the linear density dependence.

This is not to say that we substitute the formal equivalences for r and K into the logistic model and proceed as before. These are rather unwieldy. Rather, instead we proceed with the coupled differential equations for the biota and the inorganic resource, e.g., Eq. 11.8 or some equivalent form. It may be objected that substituting two coupled differential equations in place of one (the logistic equation) is also a more complicated model, and so it is. But the coupled differential equation is constrained by the more fundamental physical principle of conservation of mass rather than the *ad hoc* biological assumption of linear density dependence. It also allows us to examine simultaneously the dynamics of both the biota and the inorganic resource – this would be useful in determining how algae in a lake affect the concentrations of nutrients in the water, for example. In this sense, physics is more fundamental than biology: all biological systems must conform to conservation of matter but they do not necessarily have to conform to linear density dependence.

Resource uptake and coexistence by two consumers: a first attempt with proportional uptake

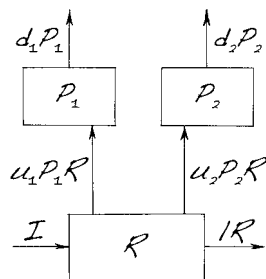
Let's turn now to the problem of coexistence of two species competing for the resource by taking it up, or "preying" upon it. What does the mass balance of the resource tell us about coexistence of the two species and the stability of the system? This approach will allow us to examine competition for a limiting resource explicitly, rather than indirectly by modeling the results of competition as reductions in growth rates in the presence of the other species, in contrast to the phenomenological Lotka–Volterra competition equations we examined in Chapter 9.

Tilman (1982) has proposed a theory of competition for resources that has gained some attention. This theory assumes that different plants compete for a single limiting resource explicitly through uptake of that resource. We will write the equations solely in terms of the density of the resource in the resource and plant pools, rather than in terms of biomass of the plants. This will allow us to dispense with the efficiency term for now, but we shall return to resource use efficiency later in the chapter.

The theory requires three equations, one each for the two plants and one for the resource. The equations are:

$$\begin{cases} \frac{dR}{dt} = I - lR - u_1 P_1 R - u_2 P_2 R = I - R(l + u_1 P_1 + u_2 P_2) \\ \frac{dP_1}{dt} = u_1 P_1 R - d_1 P_1 = P_1(u_1 R - d_1) \\ \frac{dP_2}{dt} = u_2 P_2 R - d_2 P_2 = P_2(u_2 R - d_2) \end{cases} \quad (11.12)$$

Fig. 11-2 Competition between two species for a resource in an open system.



where P_1 and P_2 are nutrient contents of plant species 1 and 2; u_1 and u_2 are uptake rates per unit plant species per unit resource per time; d_1 and d_2 are per capita death rates per time; I is input to R via exogeneous input to the ecosystem, weathering, or decay of organic matter; and l is the instantaneous export rate from R to the outside environment (Fig. 11-2). All outputs from R become either the input to a plant species or are lost from the entire system. Note that when a plant dies, the nutrients are not returned to R via litter, but are exported from the ecosystem entirely. We will take up this point later in the next chapter.

As usual, setting the differential equations equal to zero and solving simultaneously yields the equilibrium solutions:

$$R_0^* = \frac{I}{l}, P_1^* = 0, P_2^* = 0 \quad (11.13a)$$

$$R_1^* = \frac{d_1}{u_1}, P_1^* = \frac{I}{d_1} - \frac{l}{u_1}, P_2^* = 0 \quad (11.13b)$$

$$R_2^* = \frac{d_2}{u_2}, P_1^* = 0, P_2^* = \frac{I}{d_2} - \frac{l}{u_2} \quad (11.13c)$$

When both species are absent (Eq. 11.13a, the “empty resource pool” equilibrium), the equilibrium resource pool is determined by the ratio of inputs (I) to outputs (l). When one species “coexists” in monoculture with the resource (Eq. 11.13b,c), the species determines the size of the resource pool by the ratio of its death rate, d_i (which represents export from ecosystem – NOT transfer to litter) to its uptake rate, u_i . R^* is now determined solely by the traits of the plant which takes it up, not by the abiotic input–output budget (I and l). The plant is controlling R^* “from the top down”, a point to which we shall return in Chapter 13 when we consider consumers of the plants.

In addition, there is a critical value that I must exceed in order for the plant to “invade” the “empty resource pool” (Eq. 11.13a). This is a transcritical bifurcation in which Equilibria 11.13b or 11.13c exchange stabilities with the “empty resource pool” equilibrium as I exceeds this critical value. This critical I can be found by setting the nonzero P_i^* in either resource-plant equilibrium to zero and solving to find $I_{critical} > ld_i/u_i$. Whereas for Eq. 11.13a, I must simply be positive for $R_0^* > 0$, now I needs to increase above some minimum positive number (ld_i/u_i) for both R^* and

one P_i^* to be positive. If, for example, the resource is dissolved nutrient in a lake, then as input increases (due to fertilizer runoff), there is a critical value above which a plant, such as an algae or diatom, can invade. This is not unlike the sudden regime shift in lakes during eutrophication, in which there is rapid growth of a species as nutrient inputs to the lake exceed some threshold value (Carpenter 2003). Because the critical value of this input depends on the life history traits of the species (d_i and u_i), it is not possible to set a single critical input level for protecting the lake from eutrophication caused by runoff from its watershed.

Note that there is no coexistence equilibrium. For a given set of parameter values, only one species can persist, driving the other to extinction. Tilman (1982) argues intuitively that the species that has the lowest R^* (i.e., the lowest equilibrium loss rate in relation to uptake) wins by lowering R to less than the R^* for the competitor. This can also be proved as follows: Designate R_1^* as the R^* when only species 1 survives and R_2^* as the R^* when only species 2 survives. Now ask: what must be the relationship between u_1 , u_2 , d_1 , and d_2 at R_1^* such that $dP_2/dt < 0$ and therefore $P_2 \rightarrow 0$? To find this, insert $d_1/u_1 (= R_1^*)$ for R into the differential equation for P_2 :

$$\frac{dP_2}{dt} = P_2 \left(\frac{u_2}{u_1} d_1 - d_2 \right)$$

Now, $dP_2/dt < 0$ when

$$\frac{u_2}{u_1} d_1 - d_2 < 0$$

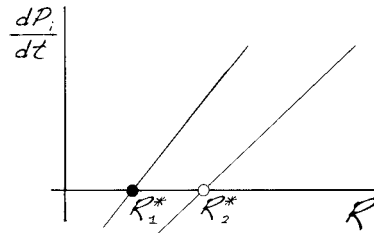
which happens when

$$\frac{d_1}{u_1} < \frac{d_2}{u_2}$$

or when $R_1^* < R_2^*$ since $R_i^* = d_i/u_i$. Therefore, if $R_1^* < R_2^*$, species 1 forms a monoculture and species 2 goes extinct. Similar arguments can be made for when species 2 survives and species 1 goes extinct.

If we plot dP_i/dt against R (Fig. 11-3), it can be seen that $dP_i/dt = 0$ when $R_i^* = d_i/u_i$. The lines are straight lines and, by virtue of the above proof, the species with the lower R_i^* will always have its growth line lie above the other. Therefore, having greater

Fig. 11-3 In Eq. 11-12, the species with the lower R^* has the greater growth rate and therefore outcompetes the other.



growth, it will acquire more and more of the resource as its biomass grows, eventually driving R into the region where the growth of the other species is negative.

Species decrease R^* by either decreasing their death rate (and hence retaining nutrients) or increasing their uptake rate (and hence increasing one of the outputs from R at the expense of the alternative uptake by the competitor). Note that mass balance requires that d and u must be correlated to maintain the same R^* : as d increases, so must u in the same proportion to maintain the same ratio at equilibrium, and vice-versa.

We can examine the stability of a species monoculture by examining the Jacobian evaluated at either Eq. 11.13b or c. This Jacobian is:

$$J = \begin{bmatrix} l \left(l - \frac{u_i I}{d_i} \right) & -d_i \\ \frac{u_i I}{d_i} - l & 0 \end{bmatrix}$$

It can be easily shown (which you should verify on a piece of scrap paper) that the trace will be negative and the determinant positive, and the solution therefore stable, when $I > l(d_i/u_i)$ or when $I > lR_i^*$. In other words, the input to R must be greater than the export from R without plant uptake, the difference being the amount of the resource which is available for plant uptake to support a positive growth of P_i^* . This implies a subtle but important point: R_i^* is *not* the availability of the resource to species i ; the availability of the resource is the difference between I and lR_i^* required to maintain a positive uptake by P_i which (at equilibrium) is balanced by the loss through death of the plant. Availability of the resource to the plant is determined by *the net flux* through R in the absence of plant uptake, not the size of R itself. If that net flux is small, then uptake is low and the size of P_i^* is small. P_i^* increases in proportion to the *difference* between I and lR_i^* . And, as we have seen, if I is below some critical value, neither plant can invade the resource pool.

Although the plant monocultures (Eq. 11.13b or c) are stable under these conditions, they are stable in a very delicate way. They are only stable to perturbations to P_i and R_i , that remove material from these pools (by harvesting the plant or “eroding” the resource) or add to these pools (by planting additional plants or fertilizing the resource) near their equilibria. Each equilibrium is not stable to invasion by the other species if the other species has a lower R_i^* . The first kind of stability is the sort that we have been investigating so far in this book. It can be termed “ecological stability” – recovery to an equilibrium point after perturbing one or more of the extant variables of the model. The second kind of stability is termed an “evolutionarily stable strategy” (Maynard Smith and Price 1973). An “evolutionarily stable strategy” (an ESS) is a configuration of species and/or resources that is stable to invasion by another species or mutant with any other set of parameter values – that is, any other set of parameter values will not survive when introduced into a system which is at an ESS. You can think of species 2 being either a separate species or a mutant variety of species 1 in which either d_2 or u_2 or both have “mutated” to yield a lower R_2^* . Each of the solutions are therefore “ecologically stable” with respect to perturbations to the equilibrium pool sizes but they are not “evolutionarily stable” to invasion by another species or new mutant.

What is happening here is that a transcritical bifurcation occurs when the R_i^* of both species are equal. To see this, examine Fig. 11-3 again. In this figure, R_i^* represents the position of the equilibrium projected onto the R axis. Suppose we begin with $R_1^* < R_2^*$. Therefore (P_1^*, R_2^*) is the stable equilibrium and (P_2^*, R_1^*) is the unstable equilibrium. Now imagine that R_2^* is decreased and allowed to approach R_1^* . Eventually they collide and as R_2^* becomes less than R_1^* , the stabilities of the equilibria are exchanged. This is precisely the defining behavior of transcritical bifurcations, as we saw in Chapter 5 for the logistic model when r passed through 0. Therefore, there is a critical point in this model, namely when $d_i/u_i = d_j/u_j$. When this critical point is passed the two equilibria with plants exchange stabilities, the one with the lower d/u being ecologically stable. There is then a shift from one monoculture to another.

Therefore, the model is stable to one kind of perturbation (harvesting or stocking or fertilizing) but not to another (species invasion). It is crucial that this difference be understood in any kind of experimental test.

Tilman (1988) suggested that two or more species can coexist at fixed point equilibria if one expands the model to include multiple resources and each species has the lowest R^* for only one resource. Thus, if the resource ratios within the biomass of each species differ and if there are multiple limiting resources, then species can coexist. Tilman and Pacala (1993) develop this line of reasoning further, but admit that this modification requires as many unique resources as coexisting species. Clearly, there are more species than material resources (some 20 or 30 essential elements plus light and water), even in a relatively depauperate temperate ecosystem, to say nothing of a species rich tropical ecosystem. One could argue that resources are spatially segregated beneath different rooting zones and could therefore be viewed as “different” resources. However, one runs the risk of pushing this argument to a tautology, since no two plants or their roots could ever occupy exactly the same space and so one can in theory generate as many resources as one needs simply requiring that different individuals occupy different spaces. The theory then becomes unfalsifiable: if the theory does not predict coexistence when in fact it is observed, simply “generate” another “resource” by spatial or temporal segregation.

Michaelis–Menten uptake of the resource by the plant

We have a problem. Using the nonmechanistic Lotka–Volterra competition equations in Chapter 9, we could get stable coexistence. In that case, the resource the species were competing for and the method of competition was not explicit. But when the resource and methods of competition were made explicit using Lotka–Volterra uptake functions, we lost coexistence and probably stability. Making things more explicit has actually led to worse model behavior. What has gone wrong?

Perhaps the problem with noncoexistence in the model lies in the Lotka–Volterra uptake functions, which are linear per capita growth functions. We have seen that the growth function of the species with lowest R^* will always lie above the other, and so that species will outcompete the other by driving down R . However, if uptake is satiated at some maximum level, as with hyperbolic or Michaelis–Menten uptake or harvesting functions, then perhaps the ability of a species to always outcompete the other will somehow be curtailed.

It is a common practice to “take an equation off the shelf” and use it. Usually this presents no problem if the proper interpretation is made of the parameters. However,

it is always good to verify that in fact the equation really can be derived from some more fundamental assumptions about the process being investigated. O'Neill et al. (1989) showed that the Michaelis–Menten function can be derived from more fundamental considerations of the supply rate of nutrients to roots and the ability of roots to take them up once they arrive. This derivation allowed O'Neill et al. (1989) to generalize the Michaelis–Menten uptake function to two nutrients, depending on further assumptions of how they interact. We will consider these extensions in Chapter 14 on stoichiometry, but for now let's examine the derivation and see what ecological conclusions we can draw from it.

O'Neill et al. (1989) begin by assuming that a producer can be in either one of two states: it can be waiting for a nutrient to arrive at an uptake site, or it can be taking up (handling) a nutrient which has already arrived (the same considerations also apply to the arrival of photons of light on a leaf and whether the chlorophyll molecule it hits is already in an excited quantum state or not). Let the probability that it is waiting be P_s and the probability that it is handling a nutrient be P_h . It is also assumed that

$$P_{s,t} + P_{h,t} = 1$$

That is, the producer does nothing else but waits for and handles nutrients. Now also assume that nutrients arrive at the root at random at a rate $v(t^{-1})$ and are taken up at a maximum rate $U_{\max}(t^{-1})$. This assumption of a maximum uptake rate is not arbitrary, but actually rests on specific biochemical reactions: uptake of a nutrient across a root cell wall occurs by means of ion-channel pumps which transport the ion across the cell wall while simultaneously maintaining a charge balance, and these ion pumps have maximum reaction rates (Taiz and Zieger 2002).

Consider now the state of the producer at time $t + \Delta t$, where Δt is a very small time interval. The probability that a producer is waiting for a molecule or ion of a nutrient at time $t + \Delta t$ is:

$$P_{s,t+\Delta t} = P_{s,t}(1 - v\Delta t) + P_{h,t}U_{\max}\Delta t$$

That is, the first term on the r.h.s is the product of the probability of waiting at t times the probability of a nutrient not arriving during Δt and so the producer is still waiting; the second term is the probability of handling at t times the maximum rate of handling during Δt and so the producer is no longer handling at the end of Δt , and therefore must be waiting for a nutrient to arrive.

The probability that the producer is handling (taking up) a molecule or ion of nutrient at time $t + \Delta t$ is:

$$P_{h,t+\Delta t} = P_{s,t}v\Delta t + P_{h,t}(1 - U_{\max}\Delta t)$$

As before, the first term on the r.h.s. represents the probability that the producer was waiting at time t times the probability that a nutrient arrived during Δt and so the producer is no longer waiting; the second term represents the probability that a producer was handling at time t times the probability that the producer did not finish handling during Δt and is still taking up nutrients at the end of Δt .

Expanding these two equations, subtracting $P_{s,t}$ or $P_{h,t}$ to obtain $\Delta P_{s,t}$ or $\Delta P_{h,t}$ and taking the limit as $\Delta t \rightarrow dt$ yields:

$$\frac{dP_{s,t}}{dt} = -vP_{s,t} + U_{\max} P_{h,t}$$

and

$$\frac{dP_{h,t}}{dt} = vP_{s,t} - U_{\max} P_{h,t}$$

Note the antisymmetry here in these two equations, which is a result of assuming that the probability of waiting and the probability of handling both add to 1.

Because

$$\frac{d}{dt}(P_{s,t} + P_{h,t}) = \frac{d}{dt}(1) = 0$$

the r.h.s. of both equations can be set equal to each other. We can also drop the t subscript since the sum of the probabilities is not time dependent, yielding:

$$-vP_s + U_{\max} P_h = vP_s - U_{\max} P_h$$

After a bit of algebra (try it!) this becomes

$$P_h = \frac{vP_s}{U_{\max}}$$

Substituting this into the equation for the sum of the two probabilities equal to unity

$$P_s + \frac{vP_s}{U_{\max}} = 1$$

yields:

$$P_s = \frac{1}{1 + \frac{v}{U_{\max}}}$$

The expected rate of uptake of a single ion or molecule of nutrient equals the rate of arrival of a limiting nutrient, v , times the probability that a user is ready to use it, P_s , or:

$$vP_s = u = \frac{v}{1 + \frac{v}{U_{\max}}} = \frac{U_{\max}}{1 + \frac{U_{\max}}{v}}$$

Resources or nutrients limit the uptake because of delays in their arrival caused by slow supply rates. The rate of arrival, v , is also the inverse of the mean delay time, τ , i.e., $v = 1/\tau$. Substitution yields:

$$u = \frac{U_{\max}}{1 + U_{\max} \tau}$$

It now remains to express τ in terms of nutrient concentration, or supply rate. The higher the nutrient concentration, the lower the delay time. Nutrient concentration is the resource density, R , since density is a concentration per unit area or per unit volume. Therefore τ is inversely proportional to R :

$$\tau = \frac{c}{R}$$

where c is the proportionality constant. Substituting into the above equation for u yields the Michaelis–Menten per capita uptake equation:

$$u = \frac{U_{\max} R}{\alpha + R} \quad (11.14)$$

where $\alpha = cU_{\max}$ and is the half-saturation constant. We have now obtained the same form of a Michaelis–Menten harvest function for uptake of an inorganic resource by a plant as we obtained in Eq. 7.12 for harvest of a predator by a prey, both beginning with the assumption that uptake or harvesting is accomplished by two mutually exclusive activities. However, in this case, we have an explicit relationship between uptake of an inorganic nutrient and its supply rate, which is determined by the nutrient density and the delay in arrival of an ion of the nutrient to the root. We are now on solid biological as well as mathematical grounds for coupling either proportional uptake or Michaelis–Menten uptake of a resource with the dynamic equation for the input–output budget of the resource pool.

Resource uptake and coexistence by two consumers: a second attempt with Michaelis–Menten uptake

Let's now incorporate the Michaelis–Menten resource uptake function into our model. In Eq. 11.12, both species had identical forms of uptake, although the parameter values have to be different if the species are to be considered distinct. But what if one species had Michaelis–Menten uptake kinetics while the other species had Lotka–Volterra uptake kinetics? Two species can be distinguished not just by different parameter values, but also by different functional forms for growth, uptake, etc. We sometimes think that a species is defined by a set of unique parameter values, but it can also be uniquely defined by having a different functional form for one of its life strategies.

Armstrong and McGehee (1980) show that letting one of the plant species have a Michaelis–Menten uptake function will allow coexistence under some conditions. I will use their reasoning in what follows. However, in their model, they assumed that the resource dynamics in the absence of the two consumers is a logistic function, but in order to stay within the spirit of our argument so far we will assume that resource dynamics in the absence of consumers is given by Eq. 11.1, i.e. with an open inorganic resource pool. The equations are:

$$\begin{cases} \frac{dR}{dt} = I - lR - \frac{U_{1\max}RP_1}{\alpha + R} - u_2P_2R = R\left(\frac{I}{R} - l - \frac{U_{1\max}P_1}{\alpha + R} - u_2P_2\right) \\ \frac{dP_1}{dt} = \frac{U_{1\max}RP_1}{\alpha + R} - d_1P_1 = P_1\left(\frac{U_{1\max}R}{\alpha + R} - d_1\right) \\ \frac{dP_2}{dt} = u_2P_2R - d_2P_2 = P_2(u_2R - d_2) \end{cases} \quad (11.15)$$

The equilibria are:

$$R_0^* = \frac{I}{l}, P_1^* = 0, P_2^* = 0 \quad (11.16a)$$

$$R_1^* = \frac{d_1\alpha}{U_{1\max} - d_1}, P_1^* = \frac{I(U_{1\max} - d_1) - ld_1\alpha}{d_1(U_{1\max} - d_1)}, P_2^* = 0 \quad (11.16b)$$

$$R_2^* = \frac{d_2}{u_2}, P_1^* = 0, P_2^* = \frac{I}{d_2} - \frac{l}{u_2} \quad (11.16c)$$

It still appears that we do not have a stable coexistence equilibrium, but as Armstrong and McGehee (1980) show, this is not exactly true. It is true that there is not a stable *fixed point* equilibrium that allows coexistence, but what actually happens is more interesting.

First, note that Eq. 11.16c is the same as Eq. 11.13c, the P_2 monoculture equilibrium for Eq. 11.12. We already know that this equilibrium is stable for perturbations to P_2^* and R_2^* but species 1 can invade if $R_1^* < R_2^*$.

What happens if $R_1^* < R_2^*$? Then we have Eq. 11.16b. The eigenvalues here are such that this could be a stable fixed point equilibrium or not, depending on parameter values. If this fixed point equilibrium is stable, then species 1 can displace species 2. But what if it isn't stable? Then, there could be a stable limit cycle between P_1 and R . We can check this by applying Kolmogorov's theorem as in Chapter 8. The per capita growth rates of P_1 and R are:

$$\begin{aligned} R : f_1 &= \frac{I}{R} - l - \frac{U_{1\max}}{\alpha + R}P_1 \\ P_1 : f_2 &= \frac{U_{1\max}R}{\alpha + R} - d_1 \end{aligned}$$

Recall that Kolmogorov's theorem gives ten conditions for there to be either a stable point or stable limit cycle. It is straightforward to show with partial derivatives of f_1 and f_2 against P_1 and R , that conditions (i) through (v) hold (try it!). Conditions (vi) and (vii) give us a bit of trouble which we will solve, and conditions (viii) through (x) will tell us something interesting. Let's go through these.

(vi) $f_1(0,0) > 0$. For this condition, if we simply plug in $(0, 0)$ for (R, P_1) in f_1 , the first term will be $I/0$, which is undefined. However, recall that this condition ensures

that the resource will grow when the densities of both the resource and the plant are near zero. This condition therefore ensures that $(0, 0)$ is unstable to small positive perturbations of R and P_1 , and so the system will grow away from the origin towards a stable point or stable limit cycle should one exist. Let's state this condition as $f_1(\varepsilon, \varepsilon) > 0$, where ε is a small amount of R and P_1 each within ε of 0. $(\varepsilon, \varepsilon)$ is now a small perturbation to R and P_1 near the origin into the first quadrant. This results in the condition:

$$\frac{I}{\varepsilon} - l - \frac{U_{1\max}\varepsilon}{\alpha + \varepsilon} > 0$$

We can now isolate the terms in ε on the r.h.s, yielding:

$$I\alpha > l\varepsilon(\alpha + \varepsilon) - l\varepsilon + U_{1\max}\varepsilon^2$$

Now taking the limit as $\varepsilon \rightarrow 0$, we obtain

$$I\alpha > 0$$

which is obviously true.

(vii) $f_1(0, A) = 0$. Condition (vii) requires that there must be a positive plant density large enough to stop the growth of the resource even when the resource is scarce (i.e., within ε of 0). This condition prevents uninhibited (unstable) growth of the resource. Kolmogorov's theorem puts this requirement on the per capita growth function, but as with condition (vi) we also run into problems with the first term being undefined when $R = 0$. However, this condition will be satisfied if $dR/dt = 0$ when $R = \varepsilon$:

$$0 = I - l\varepsilon - \frac{U_{1\max}\varepsilon}{\alpha + \varepsilon} P_1$$

and we now solve for P_1 to yield:

$$P_1 = (I - l\varepsilon) \left(\frac{\alpha + \varepsilon}{U_{1\max}\varepsilon} \right)$$

P_1 is positive when $I > l\varepsilon$, or when inputs of the resource are greater than outputs in the absence of species 1, allowing some surplus of R for species 1 to take up. So this condition holds within an arbitrarily small neighborhood ε of the origin.

$$(viii) \quad f_1(B, 0) = \frac{I}{B} - l = 0. \text{ This holds when } B = I/l, \text{ or } B = R_0^*.$$

$$(ix) \quad f_2(C, 0) = \frac{U_{1\max}C}{\alpha + C} - d_1 = 0. \text{ This holds when } C = d_1\alpha/(U_{1\max} - d_1), \\ \text{or when } C = R_1^*.$$

(x) $B > C$. This holds when $R_1^* < R_0^*$. This has to be the case for species 1 to “invade” a pure unexploited resource, otherwise if $R_0^* < R_1^*$ then species 1 will have a negative growth when its population is low during the initial stages of invasion of a system at or near R_0^* .

Therefore, in slightly modified form, all the conditions of Kolmogorov’s theorem are satisfied and there is either a stable fixed point equilibrium or stable limit cycle between P_1 and R . When there is a stable fixed point equilibrium, $R_1^* < R_2^*$ and species 1 displaces species 2, as we have seen.

But Armstrong and McGehee (1980) show that when there is a stable limit cycle, then species 1 does not exclude species 2 if the average value of R over an oscillation of period τ is greater than R_2^* . To see this, first consider that species 2 can invade the stable limit cycle of P_1 and R if and only if its average per capita growth rate over the cycle is positive. The per capita growth rate of P_2 is $u_2R - d_2$. We need to integrate over the period of the limit cycle, τ , to get the average per capita growth rate of P_2 over the cycle. In other words,

$$\frac{1}{\tau} \int_0^\tau \frac{1}{P_2} \frac{dP_2}{dt} dt = \frac{1}{\tau} \int_0^\tau (u_2R(t) - d_2) dt = -d_2 + u_2 \frac{1}{\tau} \int_0^\tau R(t) dt > 0$$

But since

$$\frac{1}{\tau} \int_0^\tau R(t) dt = \bar{R}$$

this reduces to

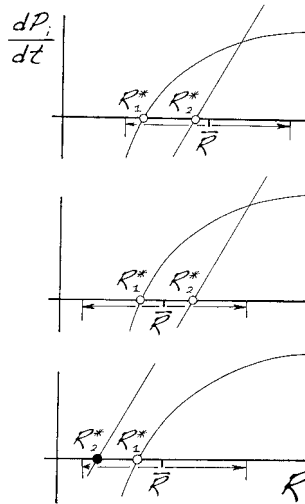
$$u_2\bar{R} - d_2 > 0 \text{ or } \bar{R} > \frac{d_2}{u_2} = R_2^*$$

Therefore, so long as there is a stable limit cycle between P_1 and R and $R_1^* < R_2^* < \bar{R}$, then the two species can coexist on one resource. But coexistence in a limit cycle comes at the expense of a stable fixed point equilibrium.

You can also see this graphically by plotting dP_i/dt against R for both species with $R_1^* < R_2^* < \bar{R}$ (Fig. 11-4). The range of variation in R during one period is symmetric around $\bar{R}$. If $\bar{R} > R_2^*$ (Fig. 11-4, top panel) then dP_2/dt will be positive for a longer period of time than it is negative, and species 2 can invade a monoculture of species 1 and persist. It is apparent that the two dP_i/dt also cross at some value $R_{\text{cross}} > R_2^*$. When R is greater than this during a limit cycle, then the growth of species 2 exceeds that of species 1 but when it is less than this then the growth of species 1 exceeds that of species 2. Therefore, each species gains a competitive edge over the other along some portion of the stable limit cycle, but neither remains at such a favorable growth rate long enough to drive the other to extinction.

But if $R_2^* > \bar{R}$ (Fig. 11-4, middle panel), then dP_2/dt will be negative for a longer period of time than it is positive. Species 2 will then go extinct and invasion will not be successful. The oscillations, however, also open an opportunity for species 2 to invade a monoculture of species 1 and subsequently displace species 1 and form

Fig. 11-4 Species 1 has a Michaelis-Menten uptake function and species 2 has a Lotka-Volterra uptake function. If there is a stable limit cycle between species 1 and R_1 , then species 2 can invade if R_2^* is less than the average size of R during a cycle (top panel). The two species can coexist in a limit cycle with R . But if R_2^* is greater than the average size of R , then it cannot invade (middle panel). However, if $R_2^* < R_1^*$, then species 2 will displace species 1 and form a stable monoculture (bottom panel).



a stable monoculture (Fig. 11-4, bottom panel). This will happen when $R_2^* < R_1^*$ but R_2^* still remains within the range of oscillation of R . Thus the oscillations allow coexistence with a stable limit cycle but they do not necessarily guarantee an evolutionarily stable equilibrium, preventing invasion of a monoculture of one species by a second.

It is because R oscillates that species 2 can coexist with species 1 when $R_1^* < R_2^* < \bar{R}$. To have oscillations and a stable limit cycle, the eigenvalues must be complex conjugates with positive real parts. The real part of the complex eigenvalues of (P_1^*, R_1^*) is:

$$\text{Re}(\lambda) = \frac{I(2U_{1\max}d_1 - U_{1\max}^2 - d_1^2) - ld_1^2\alpha}{2U_{1\max}d_1\alpha}$$

This is negative, and there is no stable limit cycle and species 2 therefore cannot invade, when the numerator is negative. Let's solve for this in terms of I , yielding

$$I_{\text{critical}} = \frac{ld_1^2\alpha}{2U_{1\max}d_1 - U_{1\max}^2 - d_1^2}$$

That is, for $\text{Re}(\lambda) < 0$ and (P_1^*, R_1^*) to be stable, I must be less than I_{critical} . If we increase inputs of the resource to the point where I_{critical} is exceeded, perhaps by fertilizing, then the equilibrium point (P_1^*, R_1^*) becomes unstable and, because Kolmogorov's Theorem is satisfied, we are left with a stable limit cycle, enabling species 2 to invade and coexist with species 1. This is exactly the same as Rosenzweig's Paradox of Enrichment, which we encountered in Chapter 8. When K for the prey in that model exceeded a critical level, the coexistence equilibrium became unstable and, again because Kolmogorov's Theorem was satisfied, we were also left with a stable limit cycle.

In this model, the resource dynamics are controlled not by a carrying capacity term, but by the mass balance of inputs and outputs. Species 1 in this model has exactly the form as the predator in the Rosenzweig–MacArthur model. Species 1 is therefore simply a predator on the resource. Increasing the input of the resource has the same consequences as increasing K in the Rosenzweig–MacArthur model: it destabilizes the fixed point equilibrium between species 1 and R through a Hopf bifurcation. This allows species 2 to invade and coexist with species 1 under suitable conditions specified above.

Once again, we see that increasing nutrient inputs to a lake (for example) above a critical value allows another species to invade and destabilize the system (in the sense of the appearance of a limit cycle instead of a fixed point equilibrium of the native species and the resource), a not uncommon consequence of eutrophication (Carpenter 2003).

We have therefore recovered all the consequences of predator–prey models with stable limit cycles as well as the possibility of coexistence equilibrium between two competitors, but we have done it without the use of K at any level of the food web. K has simply been replaced by mass balance of a limiting nutrient at the lowest, abiotic level of the food web. Therefore, the mass balance constraints of the flow of materials through a food web, required by conservation of matter, ultimately underlies the stabilities of populations, communities, and ecosystems.

Growth in relation to internal nutrient content: the cell quota and Droop's model

If we assume that production (growth of biomass) is proportional to uptake, then we can extend the Michaelis–Menten uptake equation to a production equation by rescaling $U_{\max}$ and α in terms of biomass production instead of resource uptake. If the resource limits production, then we have

$$\frac{dB}{dt} = \frac{P_{\max} R}{\alpha_p + R} \quad (11.17)$$

where $P_{\max}$ is the asymptotic rate of biomass production (dB/dt) as $R \rightarrow \infty$ and α_p is the concentration of R in which $dB/dt = 1/2 P_{\max}$.

But this equation glosses over the fact that the conversion of external nutrients into biomass is actually done in two steps: first the nutrient is taken up, then the internal concentration of the nutrient in a cell determines cell growth. Droop (1968) noticed experimentally that dB/dt is inversely related to $1/Q$, where Q is the mass of a limiting nutrient in a cell, or the cell quota:

$$\frac{dB}{dt} = a - b \left(\frac{1}{Q} \right)$$

If a = maximum growth rate ($P_{\max}$), then

$$\frac{dB}{dt} = P_{\max} - b \left(\frac{1}{Q} \right)$$

Now factor out $P_{\max}$,

$$\frac{dB}{dt} = P_{\max} \left(1 - \frac{b}{P_{\max}} \left(\frac{1}{Q} \right) \right)$$

Assume $b/P_{\max}$ equals some value z such that when $Q = z$, $dB/dt = 0$. Then $z = Q_{\min}$, or the minimum cell quota required before growth can begin. Therefore,

$$\frac{dB}{dt} = P_{\max} \left(1 - \frac{Q_{\min}}{Q} \right) = \frac{P_{\max}(Q - Q_{\min})}{Q} \quad (11.18)$$

which is a hyperbolic equation whose x -axis intersect is displaced from the origin by $Q_{\min}$ and where $P_{\max}$ is the asymptotic rate of production as $Q \rightarrow \infty$. This cell quota model, or the Droop model as it is often called, has been very useful for describing production of single celled phytoplankton in relation to some limiting resource (Droop 1968). Sterner and Elser (2002) also suggest that the Droop model can be expressed in terms of the nutrient to biomass or nutrient to carbon ratios in the cell simply by multiplying $Q_{\min}/Q$ by $1/C:1/C$:

$$\frac{dB}{dt} = P_{\max} \left(1 - \frac{Q_{\min}/C}{Q/C} \right) \quad (11.19)$$

The differential equation for Q is a simple input – output balance (i.e., uptake u of R – incorporation of R into new biomass at a per capita growth rate p):

$$\frac{dQ}{dt} = u - pQ \quad (11.20)$$

for which at steady state:

$$u = pQ \quad (11.21)$$

Because the Michaelis–Menten (Eq. 11.17) and Droop (Eq. 11.18) models are both hyperbolic equations, you may have a vague feeling that there should be some relationship between the two, and indeed there is.* We can transform the Droop model into the Michaelis–Menten model by setting $Q - Q_{\min} = Q'$. This effectively translates the Droop model to the origin. Now the Droop model becomes:

$$\frac{dB}{dt} = \frac{P_{\max} Q'}{Q_{\max} + Q'}$$

which is a Michaelis–Menten model with $\alpha = Q_{\min}$. While α could in fact numerically equal $Q_{\min}$ in real algae, there is no necessary biological or mathematical reason why it needs to. The Droop model therefore combines both the minimum cell quota with the rate of approach to $P_{\max}$ and is consequently a special case of a Michaelis–Menten model. The full model is a Michaelis–Menten model with a minimum that is independent of the half saturation constant:

*Thanks to Bruce Peckham, Dept. of Mathematics and Statistics, University of Minnesota, Duluth for suggesting this.

$$\frac{dB}{dt} = \frac{P_{\max}(Q - Q_{\min})}{\alpha + (Q - Q_{\min})} \quad (11.22)$$

If $\alpha = Q_{\min}$, this reduces to the Droop model (Eq. 11.18). We will return to this particular form of the Michaelis–Menten model later in this chapter when we examine resource use efficiency.

Burmaster (1979) showed that, at steady state, any two combinations of Michaelis–Menten resource uptake (Eq. 11.14), Michaelis–Menten biomass production (Eq. 11.17), and the Droop model (Eq. 11.18) imply the other. This means that the constraints of mass balance of resource uptake and incorporation into biomass require that if one model is true, then all three are true. To demonstrate this, we will begin with the two Michaelis–Menten equations (Eqs. 11.14 and 11.17), accept the steady state solution for Q specified by Eq. 11.21, and show that the Droop model (Eq. 11.18) follows.

To begin, note that both the Michaelis–Menten uptake and production equations are in terms of R . We can therefore solve each for R , obtaining

$$R = \frac{P\alpha_p}{P_{\max} - P} \quad \text{and} \quad R = \frac{u\alpha_u}{U_{\max} - u}$$

(to have parallel and less confusing notation, we rewrite α from the Michaelis–Menten uptake equation (Eq. 11.14) as α_u , to help distinguish it from α_p from the Michaelis–Menten biomass production equation (Eq. 11.17).) We can now equate the r.h.s. of these two equations:

$$\frac{P\alpha_p}{P_{\max} - P} = \frac{u\alpha_u}{U_{\max} - u}$$

Now assuming a steady state where the rate of uptake is exactly balanced by incorporation of R into biomass, substitute, from Eq. 11.21, pQ for u :

$$\frac{P\alpha_p}{P_{\max} - P} = \frac{PQ\alpha_u}{U_{\max} - PQ}$$

Now solve for P , which after a bit of algebra (try it!) we get:

$$P = \frac{P_{\max}\alpha_p}{\alpha_p - \alpha_u} \frac{Q - \frac{\alpha_u U_{\max}}{\alpha_p P_{\max}}}{Q}$$

This is now formally the same as the Droop model (Eq. 11.18) where

$$P'_{\max} \triangleq \frac{P_{\max}\alpha_p}{\alpha_p - \alpha_u} \quad \text{and} \quad Q_{\min} \triangleq \frac{\alpha_u U_{\max}}{\alpha_p P_{\max}}$$

Aside from the interesting mathematical techniques of showing that under steady state one model necessarily implies another, Burmaster (1979) also notes that the formal

equivalence implies that at steady state one cannot experimentally say whether the external resource concentration R or the internal cell quota Q controls growth. Both are equally valid predictors of biomass growth. In practice, under steady state conditions of phytoplankton growth, regression fits of either the Michaelis–Menten production model (Eq. 11.17) or the Droop model (Eq. 11.18) to experimental data of growth against either R or Q , respectively, will have the same r^2 .

Resource use efficiency

Resource use efficiency is the amount of biomass produced per unit resource, such as a nutrient, which is taken up. If R is a limiting nutrient such as nitrogen, virtually all of what is available in the environment is taken up by the plant, which is why the measured values of R^* tend to be very low (Wedin and Tilman 1993). If two species, each growing according to the Michaelis–Menten model, have the same resource uptake rate but different growth rates, then the one with greater growth rate (net carbon fixation) must have a proportionally greater $C/\text{nutrient}$ ratio. The $C/\text{nutrient}$ ratio is a measure of how much C is fixed in biomass per unit nutrient taken up, or the resource use efficiency. For example, the plant's C/N ratio is essentially a measure of how much biomass is produced per unit nitrogen taken up, or a measure of the plant's nitrogen use efficiency. Therefore, for a given value of R in the external environment, the plant that has the higher resource use efficiency also has the higher growth rate.

But since the Droop cell quota is not constant over a range of resource availabilities (Droop 1968), neither is the $C/\text{nutrient}$ ratio. It is therefore reasonable to ask how resource use efficiency changes along resource gradients. How the resource use efficiency of different species changes with resource availability may determine which one “wins” in competition for the resource.

Vitousek (1982, 1984) analyzed a large amount of empirical data from the literature to examine how nutrient use efficiency changes with the amount of nutrient in circulation in an ecosystem. He used the amount of nitrogen or phosphorus in litterfall as an index of the amount of nitrogen or phosphorus in circulation – essentially he assumed a steady state between N and P inputs to the soil in litter and the amounts of N and P release during decomposition which then became available to the plants. His index of resource use efficiency was the ratio of litter biomass to litter N and P content. When he plotted litter mass/ N or P content against litterfall N or P content, he found that N or P use efficiency increased as the amount of N and P in circulation decreased. This, however, has been criticized on the ground of autocorrelation: the plot of litter biomass/ N or litter biomass/ P against N or P , respectively, is a plot of $1/x$ against x (Knops et al. 1997; see also response by Vitousek 1997).

But there is a more serious objection to concluding that resource use efficiency increases monotonically as resource availability declines: if resource use efficiency continues to increase as resource availability declines, eventually resource use efficiency approaches infinity as resource availability approaches zero. This makes little biological sense. Either there must be a “stopping rule” that prevents resource use efficiency from increasing unbounded, or there must be some more fundamental relationship between productivity, resource availability, and resource use efficiency that prevents this from happening.

Let's assume that two species have Michaelis–Menten production response functions with respect to availability of a limiting nutrient, R . We have seen from

the Armstrong and McGehee model (Eq. 11.15) that for coexistence, the resource response curves must cross. Therefore, each species has greater resource use efficiency along some particular portion of a resource gradient. This is true whether the resource is measured as the standing pool of R or as resource availability, which is the flux through R that is available for uptake. Because they cross, the species with the greater R_i^* has greater growth at high levels of R than the species with the lower R_i^* , and vice-versa.

It is common for resource response curves of different species to cross at some level of the resource pool or resource availability. This results in the familiar shade tolerance classification for light, for example: so-called shade tolerant species that can tolerate low light levels (i.e., have a low R^* for light) also do not photosynthesize as fast in full sunlight (have a lower $P_{\max}$ with respect to light) as shade intolerant species which have a high R^* for light and a high $P_{\max}$ with respect to light (Larcher 1998). Along an exponentially decaying light gradient down through a canopy (see discussion of Lambert–Beers Law in Chapter 3), the two light response curves must cross at some point. This is why seedlings of shade tolerant species survive better beneath a canopy than shade intolerant species, leading to the familiar process of successional replacement of the aged individuals of shade intolerant “pioneer” species by the younger individuals of the shade tolerant “climax” species. Similar crossing of nutrient response curves have also been observed (Mitchell and Chandler 1939, Aber et al. 1979): species which have low R^* with respect to a limiting nutrient also have low maximum growth or uptake rates compared with species that have a higher R^* .

Let’s take this empirical observation of the crossing of resource response curves as given and examine what this implies for changes in resource use efficiency along resource gradients. We have seen that resource uptake, and therefore production, is positive only when $R > R^*$. This would also be true for a minimum resource flux needed to allow for positive growth of at least one unit of biomass (Sterner and Elser 2002). Bridgham et al. (1995) incorporate the minimum R^* necessary for $dB/dt > 0$ directly into the general Michaelis–Menten production response function (Eq. 11.22) as:

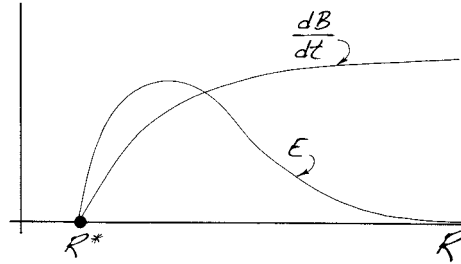
$$\frac{dB}{dt} = \frac{P_{\max}(R - R^*)}{\alpha_p + R - R^*} \quad (11.23)$$

where dB/dt is the rate of change in biomass per time (i.e., productivity) and R can be either the size of the resource pool or the flux through the resource pool (resource availability). The units of $P_{\max}$ and α need to be scaled appropriately for either measure of R . Resource use efficiency is simply the ratio of dB/dt to R , or:

$$E = \frac{dB/dt}{R} = \frac{P_{\max}(R - R^*)}{(\alpha_p + R - R^*)R} \quad (11.24)$$

If R is the flux or availability of the resource and the quantity $(R - R^*)$ is the amount taken up by the plant, then E corresponds to the nitrogen productivity concept of Ågren (1983). It is clear that efficiency equals 0 when $R = R^*$, which is also when $dB/dt = 0$. Efficiency and productivity therefore equal zero simultaneously when

Fig. 11-5 Changes in growth rate (dB/dt) and resource use efficiency (E) with increasing R . Note that both growth rate and resource use efficiency equal 0 at R^* .



$R = R^*$. Bridgham et al. (1995) now examine how efficiency changes as R increases by differentiating E with respect to R . This is a unimodal equation which is concave downward; E therefore has a maximum at a value of R which can be found by finding $\partial E/\partial R$ from Eq. 11.24, setting it equal to 0 and solving for $R_{\text{maximum } E}$, yielding:

$$R_{\text{maximum } E} = R^* + \sqrt{\alpha R^*} \quad (11.25)$$

The resource level for maximum efficiency is greater than R^* and is determined by the half saturation constant α – the smaller the half saturation constant, the more rapidly both production approaches the asymptotic level of P_{max} and the more rapidly resource use efficiency approaches its maximum. Therefore, as R increases, resource use efficiency increases to some maximum level then declines and asymptotically goes to zero as R increases (Fig. 11-5).

Pastor and Bridgham (1999) generalized these results for any monotonic and continuously differentiable function that, like the Michaelis–Menten function, meets two conditions:

- 1 There is a positive R^* such that $dB/dt = 0$.
- 2 dB/dt is monotonic and concave downward and $\lim_{R \rightarrow \infty} dB/dt = \text{Max}(dB/dt)$. This is the so-called Law of Diminishing returns, that is, as R increases further additions of R produce smaller and smaller gains in growth.

For any production function that meets these two conditions, resource use efficiency is unimodal and maximum at some value of R intermediate between R^* and as $R \rightarrow \infty$. A graphical demonstration is as follows (Fig. 11-6). For any point on a curve of production vs. resource, the efficiency of resource use is simply the ratio of production to R , or $(dB/dt)/R$. This is the slope of a line from the origin to that point. When $R = R^*$, then the line from the origin to R^* is horizontal and $E = 0$. As R increases, the slope of the line from the origin to the corresponding production increases until the line is tangent to the productivity curve, and decreases

Fig. 11-6 Resource use efficiency (the slopes of the straight lines) increases, reaches a maximum at R_2 , then declines as R increases.

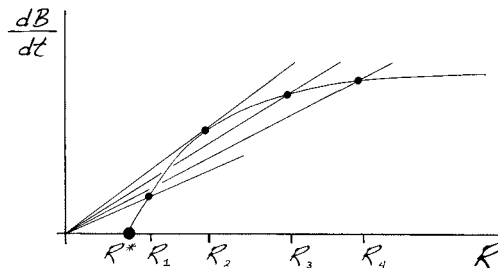
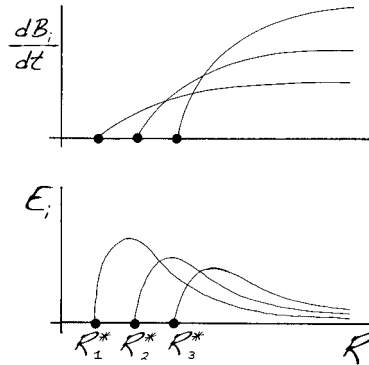


Fig. 11-7 Species replacement along a gradient of R because of changes in growth rate (dB_i/dt) and resource use efficiency (E).



with further increases of R thereafter. Because the slope is maximum when the line from the origin is tangent to the production curve, resource use efficiency is also maximum at that point and decreases with R to either side of it.

A more rigorous proof is as follows:

Condition (i) implies that at R^* , $E = (dB/dt)/R = 0$. Condition (ii) implies that $E = (dB/dt)/R \rightarrow 0$ as $R \rightarrow \infty$. For all intervening R in the interval $[R^*, \infty]$, $R > 0$ and $dB/dt > 0$, so $E > 0$. Since $E = 0$ for $R = R^*$ and since $E \rightarrow 0$ as $R \rightarrow \infty$ and, in addition, since $E > 0$ for all points in between, it must reach a maximum in the closed interval $[R^*, \infty]$. Since dB/dt is monotonic and concave downward, by Rolle's Theorem (Courant and John 1965, p. 175), at that maximum, $\partial E/\partial R = 0$. Therefore, the change in resource use efficiency with respect to R is unimodal with $E = 0$ at R^* and maximum at some positive $R < \infty$. Q.E.D.

For a series of species with progressively lower R_i^* and with Michaelis–Menten curves that cross, there is either a successive replacement of one species by another as R declines (Fig. 11-7) or else coexistence of those species with $R_i^* < \bar{R}$ if there is a stable limit cycle. Note that not only do both production and efficiency decline to 0 as $R \rightarrow R^*$, but the production and efficiency curves for any two species both cross at the same levels of R . How productivity and resource use efficiency change along resource gradients are logical consequences of one another.

Funk and Vitousek (2007) found that exotic species which have high resource use efficiency at low resource levels could invade native plant communities. This implies that these exotic invaders have a lower R^* than any of the native species. This mechanism of exotic invasion may be a logical consequence of the replacement of one species by another with lower R^* and higher resource use efficiency as resource availability declines to low levels, as seen in Fig. 11-7.

It can be seen that total ecosystem resource use efficiency – the sum total of all species resource use efficiencies – also rises to a maximum at some low level of resource availability, but at extremely low resource levels it too must decline as the lowest R^* is approached. Bridgham et al. (1995) examined nutrient uptake and production in pocosin peatlands with extremely low nutrient levels and found the steep decline in nutrient use efficiency predicted by Eq. 11.24. They also added their pocosin data to that of Vitousek (1982, 1984) and found the steep decline in efficiency at

low levels of nutrient availability for the larger dataset. Vitousek's (1982, 1984) conclusions of an unbounded increase in nutrient use efficiency with declining nutrient circulation were based on a compilation of data from the literature, the range of which apparently was not low enough to detect the decline in efficiency at low levels of nutrient circulation, perhaps because ecosystems of unusually low nutrient circulation such as pocosin peatlands were not represented.

This suggests two things: (i) We need more experimental and observational studies of productivity and nutrient use efficiency in ecosystems of extremely low nutrient availability. (ii) Theoretical conclusions drawn from compilations of data in the literature need to be treated with caution because key regions of the data field may not have been sampled or may only be sparsely sampled. Data presented in the literature is often, if not usually, gathered to test other hypotheses rather than the one being examined. Moreover, the literature, no matter how vast, represents neither a random nor a systematic sampling of the complete range of natural conditions needed to draw sound theoretical conclusions.

Summary: what have we learned?

At the lowest level of any food web are the inorganic nutrients and energy whose flux through the food web underpins biomass production. Because the flux of these resources must comply with the constraints of conservation of matter or energy, so must all the dynamics of each level of the community or food web above them. Input–output budgets of the base level limiting resource and each step in the food web above it thus constrain community and food web dynamics.

A Michaelis–Menten resource uptake function can be derived from mass balance of the resource and the assumption that uptake proceeds in two steps, each with their own probabilities: the probability of a unit of resource arriving at a reaction site (ion channel in a root cell wall or chlorophyll molecule in a leaf) and the probability that the reaction site is in a condition ready to take up the resource once it arrives. The Michaelis–Menten function can be extended to production along resource gradients. Production is also proportional to the internal content of resources within the cell (the cell quota) once they are taken up, but at steady state, Michaelis–Menten uptake and production functions imply the cell quota model. At steady state, both external resource availability and internal resource concentrations are equally valid predictors of production.

If two species have Lotka–Volterra (proportional) uptake, they cannot coexist and the species with the lowest R_i^* outcompetes the other. The two monoculture solutions are separated by a transcritical bifurcation with a critical value of $R_1^* = R_2^*$. But if one of the species with the lower equilibrium value of R (say, $R_1^* < R_2^*$) has a Michaelis–Menten uptake curve, then there is either a stable equilibrium point or a stable limit cycle between this species and R . If the average value of R during the cycle is greater than R_2^* , then the second species can invade and the two species can coexist on a single resource. This coexistence, however, comes at the expense of a stable, fixed point equilibrium and happens only because of the stable limit cycle. Explicit mass balance of an inorganic resource at the lowest level of a food web and satiation of uptake for at least one species replace any need for carrying capacity terms, which we have seen previously were necessary for stable coexistence and limit cycles. At the most basic level, mass balance of inorganic resources defines the carrying capacity of a habitat or ecosystem for each target species in a food web.

In addition, the uptake or production response curves for the two species must cross to ensure coexistence. That is, the species with the higher R^* also has a higher $U_{\max}$ or $P_{\max}$, a phenomenon which is commonly observed although the physiological basis for it is still poorly understood. The consequence of this is that resource use efficiency (production per unit resource pool or per unit resource available) is unimodal, with zero efficiency also at R^* where production is zero and maximum efficiency at some intermediate value of R . This leads naturally to species replacement as R declines, with species having lower R^* but higher maximum resource use efficiency replacing species with higher R^* but lower maximum resource use efficiency.

All these results allow us to dispense with the phenomenological linear density dependence assumptions underlying r and K in the logistic model, instead modeling population dynamics as a coupling between the biomass or resource content of the population and the input–output budget of the resource as constrained by conservation of matter. Although this coupled resource-species model may seem like a more complicated model because it replaces one differential equation (the logistic) with two coupled ones, it is on a firmer foundation because it depends on an axiomatic assumption (a conservation law) rather than an *ad hoc* assumption (linear density dependence). By extending this to additional species with suitable uptake functions, we can recover all the results of the predator–prey and competition models by considering mass balance of a limiting resource. Similar arguments can be made for mutualism, especially symbioses such as nitrogen fixation or mycorrhizal uptake of phosphorus.

This placing of population dynamics on a more fundamental physical law carries with it a subtle change in emphasis: instead of modeling the dynamics of a population as the balance between births and deaths of individuals as we have done in previous chapters, we are now modeling population dynamics as the balance between uptake of resources and the efficiency of their conversion to biomass on the one hand and losses of biomass due to metabolism and mortality on the other (Ginzberg 1998, Owen-Smith 2005). Although this is precisely the way ecosystem ecologists think about ecological systems, it may seem strange to many population biologists. However, adopting this approach has some advantages besides having a firmer foundation in physical laws. For one, it also allows us to model populations that have a large effect on the environment which we have been ignoring. For example, a herd of cattle growing to adulthood can increase its biomass by grazing while at the same time the number of individuals declines as the farmer or rancher slaughters some of them as calves for human consumption of veal: it is the increase in biomass rather than the change in the number of individuals in the herd that exerts the greatest effect on the productivity of the pasture (Owen-Smith 2005). The conservation of mass perspective also avoids the problem that many plant ecologists who worked with vegetatively propagating species such as aspen must face when trying to define an individual: is it an entire clonal structure or is it an individual stem? If we model the population dynamics of such species as changes in biomass or resource density per unit area, then we avoid such problems as to what constitutes an individual. For many lower plants such as mosses or lichens, even speaking of an individual makes little sense, and it is only by modeling the mass balance of carbon and nutrients through a moss or lichen mat that we can get a handle on their “population” dynamics.

This is not necessarily to suggest that ecologists all go out and measure the biomass or nitrogen or phosphorus content of every organism, etc., they work with and analyze population dynamics and resource consumption in terms of mass balance. Certainly, factors such as social stress in crowded situations can be the mechanisms directly underlying density-dependent birth and mortality, especially for animals such as muskrats and other rodents (Errington 1963). Even so, at the most fundamental level, ecological interactions are not so much constrained by such vague processes as density-dependence and “carrying capacity” as instead they are constrained by the mass balance of materials flowing through the community or food web from an abiotic resource which is open to the surrounding environment (Owen-Smith 2005).

Open questions and loose ends

One of the keys to stable limit cycles in the Armstrong and McGehee model, and therefore the possibility for coexistence of two or more species, is that one or more of the species must have saturated uptake or productivity rates described, for example, by Michaelis–Menten functions. The uptake or productivity functions of the different species must also cross so that there is some resource availability where each species has a higher growth rate than all other species and is therefore most efficient. Mathematically, the functions cross when R^* is correlated with $U_{\max}$ or $P_{\max}$. That such growth response functions do cross has been empirically verified many times for growth and resource uptake in response to light and in response to nutrient pool size or nutrient availability. This crossing and the asymptotic approach of growth to an upper limit leads to a natural replacement of one species by another more efficient species as the resource declines.

A theoretical basis for the crossing of the uptake or growth functions, however, is lacking. Given the importance of many of the implications of this crossing of uptake or growth functions, it would be nice to have a theoretical underpinning for it. If we could understand the physiological reasons why R^* is correlated with $U_{\max}$ or $P_{\max}$, then we can understand the physiological basis for coexistence of many species on a single resource and also for species replacement along a resource gradient. Until we have a model which shows why this is a logical consequence of some underlying process or constraint, we must take the crossing of the functions as an *ad hoc* assumption, which is not a satisfactory state of affairs.

Throughout this chapter, we have also been assuming that the plant is homogeneous with respect to nutrient concentrations. This may be an adequate assumption for single celled plants such as algae or relatively simple plants such as mosses or some of the simpler vascular plants, but it is clearly not true for large and long-lived perennial woody plants such as trees and shrubs. These have tissues whose nutrient concentrations range over nearly an order of magnitude, from wood with, for example, nitrogen concentrations of 0.5% or less to green leaves with nitrogen concentrations of 3% or higher. Thus, the efficiency of their production per unit nutrient taken up also differs. Clearly, the allocation of growth amongst different tissues in order to increase the ability of a woody plant to gather more resources by growing higher in the canopy, by making more leaves, or by making more roots has a strong effect on the plant's overall nutrient use efficiency. Each of these tissues also has a different turnover rate and therefore each has a different affect on the mean residence time of nutrients in the plant. Nutrient residence time in plants and individual plant tissues

and their relationship to plant size have strong, but as yet incompletely understood, effects on nutrient use efficiency (Berendse and Aerts 1987, Pastor and Bridgham 1999, Yuan et al. 2004). How the mass balance of resource acquisition and loss and their allocations amongst different tissues of widely different nutrient concentrations and turnover rates affects the overall efficiency of resource use is an open question that needs addressing before we can understand the full ecological meaning of efficiency.

12 Litter return, nutrient cycling, and ecosystem stability

In the previous chapter, we have seen that the mass balance of inputs and the outputs from an inorganic resource (typically a nutrient) at the base of a community or food web replaces the need for a phenomenologically defined carrying capacity to stabilize a system. But the mass balances of the Tilman and Armstrong and McGehee models considered in the previous chapter are incomplete. In particular, once the plant dies, its biomass and nutrient content are removed from the system and not returned to the resource pool as they would be in nature. The return of the nutrient contents of dead plants or dead plant parts to the resource pool occurs via litterfall and its decomposition. How does litter return and its decay affect the stability of ecosystems and the coexistence of species? These are the questions we will take up in this chapter and the next two.

In a formal sense we are now talking about an internal cycle of nutrients between different pools or compartments. This is a fundamentally different model than all the previous models which we have considered, in which materials simply flow through the system by birth or uptake followed by death or metabolism, whereupon the material exits the system (as in the predator–prey models, for example). Perhaps the uptake and litterfall that constitute a nutrient cycle are so obvious to many of us – especially those of us living in the seasonal climates of temperate and boreal regions – that we do not stop and ask why there should even be cycles. So, let's begin by asking why materials cycle in ecosystems at all.

The existence of cycles

In a very provocative book still worth reading, Morowitz (1968) first developed what he termed the material cycling theorem: In a steady state system, the flow of energy from source to sink will lead to at least one cycle in the system. Let's see how that happens.

Consider first a closed system composed of three components, A, B, and C, containing some material which is exchanged between each pair of parts in proportion to their size – that is, a linear system (Fig. 12-1). At equilibrium, the following conditions must hold:

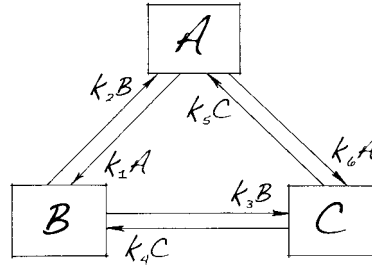
$$k_1A = k_2B$$

$$k_3B = k_4C$$

$$k_5C = k_6A$$

And so there is no net flow around the system. But if the system is closed, then the total amount of material is constant, or

Fig. 12-1 Cycling of material in a system of three components, redrawn from Morowitz (1968).



$$A + B + C = M$$

The constant amount of material in a closed system imposes limits on their distribution: if nutrient pools in two of the three above compartments are determined, then the third is automatically fixed (DeAngelis et al. 1989). This allows us to solve for the equilibrium values as:

$$A^* = \frac{k_2 k_4}{k'} M$$

$$B^* = \frac{k_1 k_4}{k'} M$$

$$C^* = \frac{k_1 k_3}{k'} M$$

$$k' = k_2 k_4 + k_1 k_4 + k_1 k_3$$

This can be summarized by stating that at equilibrium each compartment is some fraction f_i of the total system material content (M), the fraction being represented by the products and quotients of the transfer coefficients, k_i . The material or nutrients in each compartment also have some probability of being transferred to another component j , the transition probability being t_{ij} . Therefore at equilibrium, the following must hold:

$$f_i t_{ij} = f_j t_{ji}$$

for all i and j .

Now imagine that this system is irradiated with electromagnetic radiation such that there is a net absorption of the radiation by at least one component of the system (this is what happens during photosynthesis). The electromagnetic radiation that is absorbed will induce additional transitions between components as well as the previous radiation-independent transitions. Therefore, there is a new set of fractions and transition probabilities, f'_i and $t'_{i,j}$ which include both the radiation-independent and radiation-dependent transitions. Now, using our old friend proof by contradiction from Chapter 2, it can be shown that there are some pairs of components i and j for which the following does not hold:

$$f'_i t'_{ij} = f'_j t'_{ji}$$

Here's how. First assume that this does hold for all i and j . Then the absorption of radiation leading to a transition from i to j must be balanced by a reverse transition from j to i caused by the emission of a photon of radiation at the same energy. This implies that there is no net absorption of radiation by the system and so contradicts our original assumption that radiation is absorbed by the system. Therefore, there are some transitions for which the above does not hold, or

$$f'_i t'_{ij} - f'_j t'_{ji} \neq 0$$

This means that for some pathways there is a net flux from i to j or from j to i . Eventually, the system must come to a new equilibrium in which the fractions are constant or time-independent. Therefore,

$$\frac{df'_i}{dt} = 0 = \sum_j (f'_i t'_{ij} - f'_j t'_{ji})$$

But since some $f'_i t'_{ij} - f'_j t'_{ji}$ are nonzero, then others must also be and of reverse sign in order to compensate for the derivative to be equal zero. Therefore, the material must leave a component by one path and return by another. Formally, this constitutes a cycle.

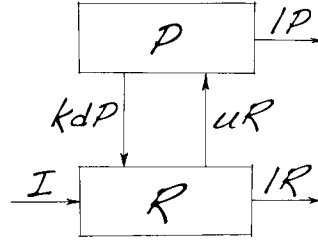
Therefore in a system closed to material inputs and outputs, the absorption of photons (energy) by one component necessarily leads to at least one cycle around several components. The absorption of photons is, of course, what photosynthesis does. Eventually this electromagnetic energy must be reradiated, otherwise we have an infinite sink whose temperature will rise exponentially. Therefore, the flux of electromagnetic energy through an open system causes material within the system to cycle.

It is well to point out that the only natural system which is closed to material cycles is the globe itself. If we think of the three components as the carbon contents of living material, dead material (including fossil fuels), and the atmosphere, then it is absorption of light radiation by photosynthesis and release as infrared radiation during respiration and decomposition on a global scale that causes carbon to cycle. But, since the beginning of the Industrial Revolution we are also greatly increasing the probability of a transition of carbon from fossil fuels to the atmosphere. This alters the heat balance of the atmosphere since carbon dioxide absorbs infrared radiation. Since the rates of virtually all of these biological and geochemical transitions also depend on temperature, we are also altering the fluxes of carbon and its equilibrium distributions amongst living material, dead material, and the atmosphere. Therefore, Morowitz's (1968) material cycling theorem, as abstract as it seems at first, has important global policy implications.

Litter return and stability: single species systems

All ecosystems at smaller spatial scales than the globe are open systems: that is, there are inputs and outputs to and from at least one component. Plants take up nutrients from this compartment and incorporate them in biomass. Upon death of the plant or plant part, the return of nutrients in the dead plant material to the inorganic pool through decay of litter is also a major pathway of nutrient flux. To model nutrient recycling through litter return, we must return some portion of the nutrient contained in biomass to the resource pool through litterfall and its decay at rate k (Fig. 12-2). For these simplest of models, we will assume that P represents the amount

Fig. 12-2 An open linear system with litter return from the plant to the resource, modeled by Eq. 12.1.



of the nutrient contained in the total organic matter in the system. Some fraction d of P is in dead biomass while the remaining fraction is in live biomass. R represents the pool of inorganic nutrient from which some fraction u is available to and taken up by P and to which P returns nutrients in inorganic form (we are assuming that the microbes decomposing the dead portion of P take up nutrients as well as the live plants; microbial uptake of nutrients is known as immobilization and we will explore this process further in Chapter 14). Some fraction k of dP is returned to R by decomposition of the dead portion of P while another portion, lP , is either consumed by herbivores and removed from the immediate plant-resource system, is exported in streamflow if we are considering lake or reach of stream, or is simply “blowing in the wind.”* Assume that the proportion of P which is exported is the same as the proportion of R that is also exported (this would especially be the case for a chemostat or a reach of a stream). This model is a simple donor-controlled model (all fluxes are controlled by the size and turnover of the donor rather than the recipient compartments), and all fluxes are linear (proportional to compartment size).

The equations are:

$$\begin{cases} \frac{dR}{dt} = I + kdP - uR - lR \\ \frac{dP}{dt} = uR - kdP - lP \end{cases} \quad (12.1)$$

At equilibrium,

$$R^* = \frac{I(kd + l)}{l(u + kd + l)}, \quad P^* = \frac{lu}{l(u + kd + l)} \quad (12.2)$$

The Jacobian is:

$$J = \begin{bmatrix} -(u + l) & kd \\ u & -(d + l) \end{bmatrix}$$

It is obvious that $\text{tr}(J) < 0$. The determinant of J is, after subtracting the appropriate cross products and doing a bit of algebra (try it!):

$$\det(J) = ud(1 - k) + l(d + u + l)$$

*From the Minnesota native, poet, and environmentalist, Bob Dylan.

which, since $k < 1$, is obviously greater than zero for all values of R and P . Therefore, adding return of the nutrient to R imposes global stability on the nutrient cycle in an open system. Further examination and generalization of how material recycling underlies the stability of both open and closed linear systems are given by Webster et al. (1975; see also Harwell et al. 1977 for further analysis of this paper) and Nisbet and Gurney (1976), respectively. Nisbet and Gurney (1976) in particular show that increasing the number of pathways of litter return from higher trophic levels increases the stability of a linear, donor-controlled system. This is in contrast to the findings of May (1975b) that increasing the number of random links in food webs without litter return decreases the chances for stability. Adding the mass balances constraints of litter return seems to be a very easy way to stabilize food web models.

What if we make uptake a Lotka–Volterra function of R and P ? Now the uptake of the nutrient is controlled both by the donor compartment (R) as well as the recipient compartment (P). Then the model becomes:

$$\begin{cases} \frac{dR}{dt} = I + kdP - uPR - lR \\ \frac{dP}{dt} = uPR - kdP - lP \end{cases} \quad (12.3)$$

The nullclines for P are the line $P = 0$ and the line

$$R = \frac{kd + l}{u}$$

Because of the constant input to R , $R = 0$ is not a nullcline. R has a single nullcline which is

$$P = \frac{lR - I}{kd - uR}$$

This nullcline is a hyperbolic function. One branch intersects the R axis at the point $R = I/l$ and is asymptotic to the line $R = kd/u$, whereupon $P \rightarrow +\infty$. The other branch is asymptotic to the same line from $-\infty$ and intersects the P axis at the point $P = -I/kd$. The positive solutions for R are therefore bounded between $R = kd/u$ and $R = I/l$ (Fig. 12-3).

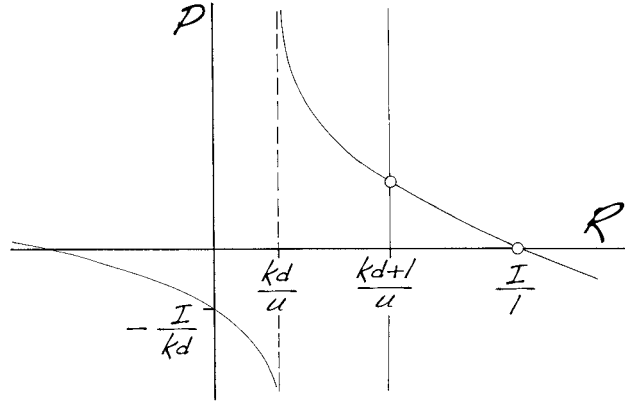
Notice that as $l \rightarrow 0$ and the system “leaks” less, the P nullcline $(kd + l)/u \rightarrow kd/u$ and $P^* \rightarrow \infty$. This is because inputs to the system do not leave and are incorporated into P .

There are two equilibrium solutions:

$$R_0^* = \frac{I}{l}, P_0^* = 0 \quad (12.4a)$$

$$R_{p+}^* = \frac{kd + l}{u}, P_{+}^* = \frac{Iu - lkd - l^2}{ul} \quad (12.4b)$$

Fig. 12-3 Nullclines and equilibria for a system with Lotka–Volterra uptake kinetics and litter return (Eq. 12.3).



(the subscript P_+ indicates the equilibrium with a positive P equilibrium). Note that increasing the uptake of nutrient into P (increasing u) also decreases $R_{P_+}^*$, thereby decreasing leaching loss from R since export via leaching is proportional to R . This is a mathematical confirmation of the “biomass increment hypothesis” of Vitousek and Reiners (1975), which states that leaching losses are inversely proportional to the rate of sequestration of nutrients into live and dead biomass.

The Jacobian at the second equilibrium with the plant present (Eq. 12.4b) is

$$J = \begin{bmatrix} -\frac{Iu - lkd - l^2}{l} - l & -l \\ \frac{Iu - lkd - l^2}{l} & 0 \end{bmatrix} \quad (12.5)$$

The trace is negative and the determinant is simultaneously positive (and the second equilibrium is stable) only when $Iu - lkd - l^2 > 0$. Since we have redirected some fraction k back to R , let's solve this inequality for k to find a critical value for decay rate required for stability in terms of the other parameters. We then get:

$$k < k_{\text{critical}} = \frac{Iu - l^2}{ld}$$

If you substitute k_{critical} for k into either P_+^* or $R_{P_+}^*$ for Eq. 12.4b, you will see that you get back Eq. 12.4a, $P_0^* = 0$, $R_0^* = I/l$. Therefore, at $k = k_{\text{critical}}$, the equilibrium with the plant present (Eq. 12.4b) becomes unstable and collapses to the empty resource pool (Eq. 12.4a), which is now stable when $k > k_{\text{critical}}$.

Exercise 12.1

Verify that the first equilibrium $(0, I/l)$ is stable when $k > k_{\text{critical}}$ by calculating the Jacobian at the first equilibrium and showing that the trace is negative and the determinant is positive only when $k > k_{\text{critical}}$.

As k increases to k_{critical} , the P nullcline ($R = (kd + l)/u$) moves to the right and approaches $R_0^* = I/l$. This causes the equilibrium with the plant present (Eq. 12.4b)

to slide down the R nullcline and approach the empty resource pool equilibrium (Eq. 12.4a). You can see that the two collide when $(kd + l)/u = I/l$. If you solve this for k , you will recover the expression we just derived for $k_{critical}$. Therefore at $k = k_{critical}$, the two equilibria collide and switch stabilities. We therefore have a transcritical bifurcation separating the two equilibria. We shall see several more examples of transcritical bifurcations separating two or more solutions in nutrient recycling models. The bifurcations (and therefore the stability of the solutions) will depend on the mass balance of nutrient transfers within the system as well as the input–output budget of the system as a whole.

The equilibrium with the plant present can therefore remain stable if the decay rate of the dead biomass k is less than $k_{critical}$. The decay rate of litter and the constraints of conservation of matter therefore determine the stability of this system.

There are some similarities but important differences between the nullclines and stability of this model compared with those of the Lotka–Volterra predator–prey model with logistic growth for the prey which we examined in Chapter 8. The plant here is a sort of “predator” on the resource and indeed the uptake function here is formally equivalent to the predator harvesting function. And just as the predator’s nullcline depends only on the prey in the Lotka–Volterra models, the plant’s nullcline also depends only on R .

But in contrast to predator–prey models, when the plant dies its nutrient content is returned to the resource pool which it “preyed” upon. There is also a constant input to R which is independent of the size of R , unlike the input to the prey population from births, which depend on the density of the prey population. The constant input to R and recycling of nutrients in this model causes the R nullcline not to intersect the P axis but instead to asymptotically approach a vertical line as exports go to zero.

The intersection of the two nullclines in the first quadrant of this model appears to resemble the similar intersection of the predator and prey nullclines of the Lotka–Volterra model (but we shall see in a moment that there are important differences in the flow of the vector fields near the equilibria). In this model, the mass balance of inputs (I) and outputs (l) plays the same role in ensuring stability as did K in the predator–prey model. In both models, if the vertical nullclines for the plant or the predator are moved to the right of I/l or K , respectively, then the plant or the predator go extinct and only R or the prey exist at equilibrium values of I/l or K , respectively.

The divergence of the vector field

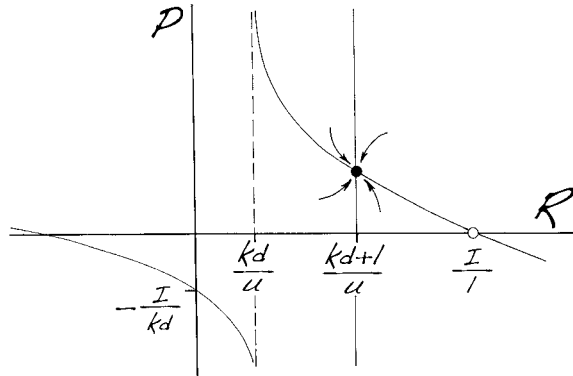
$$\text{div}(\mathbf{v}) = -\frac{Iu - lkd - l^2}{l} - l$$

is negative when $k < k_{critical}$. Likewise, the curl

$$\text{curl}(\mathbf{v}) = \frac{Iu - lkd - l^2}{l} + l$$

is positive when $k < k_{critical}$. There is therefore, on average, a counterclockwise flow of trajectories near this equilibrium (Fig. 12-4). However, unlike the predator–prey model, the eigenvalues are not complex but are both real:

Fig. 12-4 Nullclines, equilibria and their stability, and the vector field for the system with Lotka–Volterra uptake kinetics and litter return (Eq. 12.3).



$$\lambda_1 = -l, \lambda_2 = -\frac{lu - lkd - l^2}{l} \quad (12.6)$$

Again, λ_2 is negative only when $k < k_{\text{critical}}$. The positive curl causes the trajectories near this equilibrium to “curve” in to it counterclockwise, although the trajectories do not spiral around over 2π radians as they would if the eigenvalues were complex (Fig. 12-4). Therefore, although there are some similarities between this nutrient recycling model and the predator–prey model with logistic growth of the prey, the trajectories of the vector fields differ. There is no stable limit cycle in this model analogous to that in the predator–prey model.

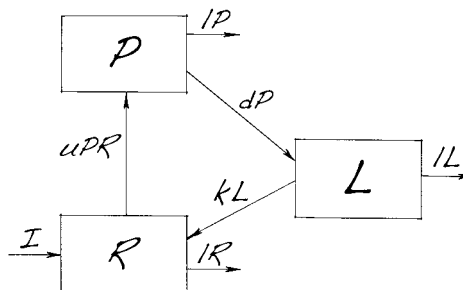
Exercise 12.2

We can separate out the dead material from P into a separate litter compartment, L . Uptake of R into P therefore depends only on uptake by live plants, and not by microbes as well (Fig. 12-5). Write the equations for this three compartment model and analyze its equilibria and their stability. Hint: If you add the three equations together, you will find that the derivative for the total ecosystem nutrient content ($R + P + L$) is:

$$\frac{d(R + P + L)}{dt} = I - l(R + P + L)$$

at equilibrium, $R^* + P^* + L^* = I/l$. You can now solve this for the equilibrium value of either R , P , or L in terms of the others, substitute this expression for the chosen variable into the differential equations for the other two compartments, and analyze the model's equilibria and their stability as a two-compartment model. This is another method of reducing the dimensions of a model to simplify its analysis.

Fig. 12-5 A three compartment nutrient cycling model with an inorganic nutrient pool (R), live plants (P), and dead litter (L).



Litter return and transient responses to perturbations

Feedbacks inherent in the nutrient cycle, such as in nutrient uptake and return in litter, can initially amplify perturbations near equilibria, even though the perturbation eventually decays and the system asymptotically converges to the equilibrium as $t \rightarrow \infty$ (Neubert and Casswell 1997). This transient amplification of a perturbation can be quite large and last for a long time, even though the real parts of all eigenvalues are negative and the system is stable in the long run (Neubert and Casswell 1997). Because the transient divergent movement can be large and long lasting, it can have important consequences. It is therefore of practical importance to understand these amplified transient behaviors of perturbations, especially for managers who usually do not have the luxury to wait as $t \rightarrow \infty$.

Since the eigenvalues only determine the long-term behavior of the system, they may not tell us anything about these transient behaviors. Moreover, two matrices can have the same eigenvalues but one can asymptotically approach equilibrium monotonically while another initially moves away from its equilibrium and then converges to it. Neubert and Casswell (1997) give the following example of two matrices:

$$J_1 = \begin{bmatrix} -1 & 1 \\ 0 & -2 \end{bmatrix}, \quad J_2 = \begin{bmatrix} -1 & 10 \\ 0 & -2 \end{bmatrix}$$

which have the same eigenvalues, $\lambda_1 = -1$, $\lambda_2 = -2$. Imagine these matrices are Jacobians of a system:

$$\frac{dx}{dt} = Jx, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which is precisely the linearization of all the two-component models we have been considering. Suppose we now introduce a perturbation away from the equilibrium x_1^* , $x_2^* = 0$ of each of the size:

$$x(0) = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

The distance of the perturbation from equilibrium at any point in time is simply the Euclidean distance:

$$\|x(t)\| = \sqrt{x_1^2(t) + x_2^2(t)}$$

For the size of our initial perturbation we can see that $\|x(t)\| = 1$. If you multiply J_1 or J_2 times $x(0) = [1/\sqrt{2}, 1/\sqrt{2}]^T$ (where the superscript T indicates transpose rows and columns), you will see that while the two systems eventually converge asymptotically to their equilibrium at the same rate (determined as usual by the dominant eigenvalue), the perturbation for the system with J_2 initially grows before shrinking while that of the system with J_1 shrinks monotonically. Therefore, while the

eigenvalues tell us the long-term behavior of a perturbation near equilibrium and thus the asymptotic (Lyapunov) stability of a system, they say nothing about the short-term transient behavior of a perturbation.

Exercise 12.3

- 1 Verify the eigenvalues of $\mathbf{J}_1$ and $\mathbf{J}_2$ and the trajectories given above. The trajectories are obtained by multiplying the perturbation $\mathbf{x}(0)$ by the Jacobians to get $\mathbf{x}(1)$, then multiplying $\mathbf{x}(1)$ by the Jacobian to get $\mathbf{x}(2)$, etc. This can easily be done with MatLab, Mathematica, or even with a spreadsheet such as Excel.
- 2 Neubert and Casswell (1997) also show that there are matrices that have the same complex eigenvalues but different transient responses, for example:

$$\mathbf{J}_3 = \begin{bmatrix} -1 & -12 \\ 0.75 & -2 \end{bmatrix}, \quad \mathbf{J}_4 = \begin{bmatrix} -1 & -4 \\ 2.25 & -2 \end{bmatrix}$$

Find the eigenvalues for each of these and plot the trajectories for an initial perturbation of $\mathbf{x}(0) = [1/\sqrt{2}, 1/\sqrt{2}]^T$ as above.

To determine the transient behavior of a system after perturbation, and especially whether it initially is amplified by feedbacks and therefore moves away from the equilibrium, Neubert and Casswell (1997) introduce the concept of reactivity. A system is stable but reactive if it initially amplifies a perturbation, but which eventually converges asymptotically back to the equilibrium. Neubert and Casswell (1997) begin by writing the growth rate of the Euclidean distance of a perturbation from equilibrium as:

$$\frac{d\|\mathbf{x}\|}{dt} = \frac{d\sqrt{\mathbf{x}^T \mathbf{x}}}{dt}$$

This can be expanded as follows:

$$\frac{d\|\mathbf{x}\|}{dt} = \frac{\mathbf{x}^T (d\mathbf{x}/dt) + (d\mathbf{x}/dt)^T \mathbf{x}}{2\|\mathbf{x}\|} = \frac{\mathbf{x}^T (\mathbf{J} + \mathbf{J}^T) \mathbf{x}}{2\|\mathbf{x}\|}$$

where $\mathbf{J}$ is the Jacobian matrix. It is obvious that the dynamics of the perturbation are governed by the term $(\mathbf{J} + \mathbf{J}^T)/2$, which is another matrix called the *Hermitian part* of $\mathbf{J}$, or $\mathbf{H}(\mathbf{J})$. The reactivity is simply the dominant eigenvalue of $\mathbf{H}(\mathbf{J})$. Assuming that the eigenvalues of $\mathbf{J}$ have negative real parts, if the dominant eigenvalue of $\mathbf{H}(\mathbf{J})$ is positive, then a perturbation will initially grow before converging asymptotically back to equilibrium, but if the dominant eigenvalue of $\mathbf{H}(\mathbf{J})$ is negative, then the perturbation will converge monotonically back to equilibrium.

If we return to the Jacobian described by Eq. (12.5), we can easily calculate $\mathbf{H}(\mathbf{J})$, for which the dominant eigenvalue is:

$$\lambda_1 = \frac{l^2kd - l u l + \sqrt{2l^2k^2d^2 - 4l u l k d + 2l^2u^2 + 3l^3k d + 2l^4 + 2l^3 - 3l u l^2 + l^2k d}}{2l} \quad (12.7)$$

When the numerator is positive, then the system is reactive and perturbations are initially amplified by the flux and circulation of nutrients. Now, this is a rather formidable looking beast, but we can get lucky and it turns out that by setting the numerator equal to 0 and solving for one parameter in terms of the others, many of the terms cancel out. This is easily done using software such as Matlab or Mathematica. Let's find a critical value for the input, I , in terms of all other parameters that separates a stable but reactive system from one which is stable but nonreactive. Perturbations to this system will be initially amplified when:

$$I > I_{crit} = \frac{l(2l + kd)}{u}$$

A large input of nutrients will therefore cause a perturbation to grow before it decays, and the larger the input the faster the growth and the greater the magnitude of growth. Thus, fertilizer input to a lake from runoff could result in very large departures from equilibrium even if the system remains stable and eventually converges asymptotically to equilibrium. Similarly, we can also solve for a critical decay rate that also determines when perturbations will be amplified:

$$k > k_{crit} = \frac{lu - 2l^2}{dl}$$

A large value for k , or a fast decay rate of dead biomass, will also increase the chance that a perturbation will initially grow. Similar calculations also show that a high uptake rate also has the same effect.

A high input rate, a high uptake rate, or a fast decay rate each increase the amount of a nutrient cycling through the system. In addition, a high input or uptake rate or a fast decay rate increase productivity because the flux of nutrients into the resource pool (and therefore available for uptake into biomass) is thereby magnified. Therefore, perturbations to productive systems are more likely to be initially amplified compared with perturbations to unproductive systems, even though both systems may be asymptotically stable to perturbations in the long run.

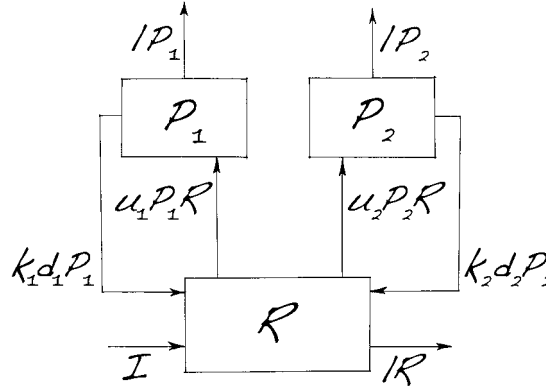
Systems in which a perturbation is initially amplified but then return to equilibrium are said to be *resilient* – as we have seen, systems with rapid uptake and decay rates are prone to be resilient. Systems in which the initial perturbation is damped monotonically are said to be *resistant*.

Litter return and coexistence of two species

In the previous chapter, we saw that two species could not stably coexist if they each have Lotka–Volterra uptake forms. But in that model, the nutrient content of a species was exported from the system upon death. What if we allow the nutrient contents of both species to be recycled into R after they die (Fig. 12-6)? This gives the other species another chance at getting the nutrient. Will litter return allow stable coexistence with Lotka–Volterra uptake forms?

To examine this, we introduce equations for two species, each with their own characteristic uptake, death, and litter decay rates:

Fig. 12-6 A two-species nutrient cycling model with litter return modeled by Eq. 12.8.



$$\begin{cases} \frac{dR}{dt} = I + k_1 d_1 P_1 + k_2 d_2 P_2 - u_1 P_1 R - u_2 P_2 R - lR \\ \frac{dP_1}{dt} = u_1 P_1 R - k_1 d_1 P_1 - lP_1 \\ \frac{dP_2}{dt} = u_2 P_2 R - k_2 d_2 P_2 - lP_2 \end{cases} \quad (12.8)$$

The equilibria are:

$$R_0^* = \frac{I}{l}, \quad P_1^* = P_2^* = 0 \quad (12.9a)$$

$$R_1^* = \frac{k_1 d_1 + l}{u_1}, \quad P_1^* = \frac{Iu_1 - lk_1 d_1 - l^2}{u_1 l}, \quad P_2^* = 0 \quad (12.9b)$$

$$R_2^* = \frac{k_2 d_2 + l}{u_2}, \quad P_1^* = 0, \quad P_2^* = \frac{Iu_2 - lk_2 d_2 - l^2}{u_2 l} \quad (12.9c)$$

As with the Tilman model discussed in the previous chapter, we still don't have a fixed-point coexistence equilibrium. This is rather surprising, since recirculation of the nutrient should allow both species to "share" the nutrient.

Moreover, the decay rate puts strict limits on the stability of either monoculture. To see this, let's look at the eigenvalues of the Jacobian evaluated at the P_1 monoculture (Eq. 12.9b). Two of the eigenvalues are:

$$\lambda_1 = -l, \quad \lambda_2 = \frac{l^2 + lk_1 d_1 - Iu_1}{l}$$

Note that when the numerator of λ_2 is negative, P_1^* is automatically positive and stable. We can solve λ_2 for k_1 as before to obtain an analogous $k_{1,critical}$ that separates the P_1 monoculture equilibrium (Eq. 12.9b) from the empty resource equilibrium (Eq. 12.9a) by a transcritical bifurcation. The third eigenvalue is

$$\lambda_3 = \frac{k_2 d_2 u_1 - k_1 d_1 u_2 + l(u_1 - u_2)}{u_1}$$

The sign of this eigenvalue separates the P_1 monoculture (Eq. 12.9b) from the P_2 monoculture (Eq. 12.9c). The sign of λ_3 depends on the numerator. Let's solve for a critical value of k_1 that separates negative values of λ_3 (and hence stability for the P_1 monoculture) from positive values by setting $\lambda_3 = 0$ and solving for k_1 , yielding:

$$k_1 < k_{1,critical} = \frac{k_2 d_2 u_1 + l(u_1 - u_2)}{d_1 u_2}$$

For stability of the P_1 monoculture, k_1 must be less than both critical values. Species 1 can achieve this in several ways: (i) by keeping k_1 , its litter decay rate, as small as possible; (ii) by keeping d_1 , the proportion of species 1 organic matter in litter, as small as possible; and/or (iii) by keeping the difference between its uptake rate, u_1 , and that of species 2 as large as possible. All of these conditions result in nutrients being sequestered in live or dead biomass of species 1 and separate a stable P_1 monoculture (Eq. 12.9b) from an unstable P_2 monoculture (Eq. 12.9c). But by symmetry, we can make the same arguments for the stability of the P_2 monoculture. Therefore, the two monoculture solutions are also separated by a transcritical bifurcation which depends on critical combinations of the two species' uptake rates, proportion in dead organic matter, and decay rates.

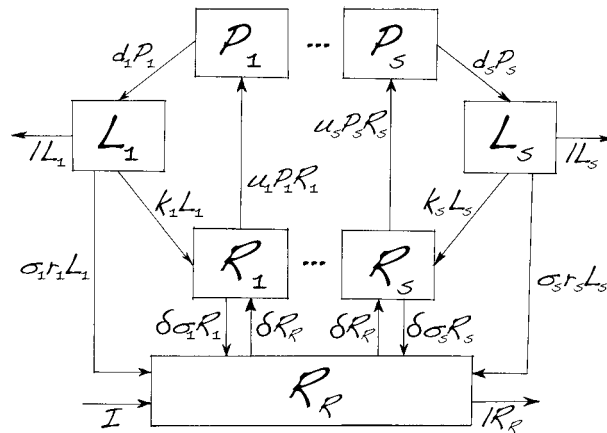
Therefore, to maintain a stable monoculture, there are critical values not only for uptake and death rates of a species but also for its decay rate. The decay rate of a species' litter is as important a life history trait as its uptake rates or its death rates. As before with a single species model, the dead material (litter) is still a part of the population – it is merely detached from the live part of the population and subject to different rules. Only now with a two-species model, the decay rate of a species' litter also determines the stability of the species' monoculture against invasion by a competitor. A species can be a better competitor not only by having greater nutrient uptake rates, which is the usual way we think of competition, but also by having slow decay rates for its dead biomass.

You can think of the three equilibria as existing in a state space of axes R^* , P_1^* , and P_2^* . These three equilibria are analogous to three ecosystems on a natural landscape: (i) an ecosystem consisting only of R with an input and an output, or a "lake" empty of organisms; (ii) an ecosystem which is a monoculture of P_1 and R ; (iii) an ecosystem of a monoculture of P_2 and R . The transcritical bifurcations are functions separating the regions of the state space occupied by the three equilibrium solutions, or ecosystems, at stability. These transcritical bifurcations depend in turn on parameter values. If these parameter values change as one moves across a landscape (because of climatic or other gradients that affect them, for example), then the bifurcations represent real boundaries between different ecosystems.

Dividing the resource pool and species coexistence

The fact that a model of explicit resource competition which also includes litter return does not allow for stable coexistence of two or more species is surprising for two reasons. First, allowing the resource to recycle through the available pool should allow

Fig. 12-7 A nutrient cycling model with S species, litter return, and partial segregation of the resource pool, redrawn from Loreau (1998).



each species to have a “second chance” at obtaining some of it and thereby avoid extinction from a superior competitor. Second, uptake of nutrients by primary producers and their return through litter and its decay accounts for greater than 90% of the flux of nutrients through most ecosystems, the amount flowing to higher trophic levels being no more than, and usually much less than, 10% of nutrients contained in plant biomass. Is the coexistence of species independent of the mass balance of nutrients cycling through an ecosystem? Clearly this is not the case, as numerous field studies have shown. One way to have stability is, of course, to have the uptake curves be Michaelis–Menten functions which cross, as we saw in the previous chapter. That model, the Armstrong and McGehee model, was able to produce coexistence with limit cycles but without litter return. Is the enormous flux of nutrients in litter and through soil and sediments not as important as the balancing flux in uptake? This asymmetry in the importance of uptake vs. the balance required in litter return and decay seems implausible. Alternatively, perhaps the underlying assumption of a homogeneously mixed resource pool needs to be relaxed to depict the proper mass balance of resource fluxes.

Loreau (1998) presents a very elegant solution to this problem in a model which is an elaboration of an earlier model by Huston and DeAngelis (1994). Loreau separates the nutrient resource into a regional, well-mixed pool (R_R) interacting with small local pools (R_i), one for each of S species, through diffusion from R_R to R_i at a diffusion rate δ (Fig. 12-7). The transfer of nutrients from live plant species (P_i) to their litter/detritus (L_i) and then through decay back to the local and regional resource pools (at rates k_i and r_i , respectively) are also made explicit. Each species i of S species acquires its nutrients from the local nutrient resource, analogous to local rooting zones, through Lotka–Volterra uptake. Loreau also separates the local resource pools into two components: the concentration of the resource per unit volume and the volume of space (σ_i) containing each local resource pool, analogous to a local rooting zone. Nutrients are exported from the regional pool and the detritus at proportional rate l as with the previous models. The mass balance of nutrient fluxes is described by the following set of coupled differential equations:

$$\left\{ \begin{array}{l} \frac{dR_R}{dt} = I + \sum_{i=1}^S \sigma_i r_i L_i - \delta \sum_{i=1}^S \sigma_i (R_R - R_i) - l R_R \\ \frac{dR_i}{dt} = \delta (R_R - R_i) + k_i L_i - u_i R_i P_i \\ \frac{dP_i}{dt} = u_i R_i P_i - d_i P_i \\ \frac{dL_i}{dt} = d_i P_i - k_i L_i - r_i L_i - l L_i \end{array} \right. \quad (12.10)$$

The equilibrium solution for the regional resource pool is

$$R_R^* = \frac{I + \delta \mu S \bar{\sigma}_i \bar{R}_i^*}{1 + \delta \mu S \bar{\sigma}_i} \quad (12.11)$$

where $\bar{\sigma}_i$ and $\bar{R}_i^*$ are the average rooting volumes and equilibrium local resource pools across all of i species, respectively and $\mu = 1/(\bar{r}_i + l)$. An interesting consequence of this equilibrium is that stable coexistence of each species with all others requires

$$I + \delta \mu S \bar{\sigma}_i (\bar{R}_i^* - R_i^*) - R_i^* > 0$$

The first two terms represent input to the ecosystem (I) plus nutrient input to the regional pool from all local resource pools, which is proportional to $\delta S(\bar{R}_i^* - R_i^*)$. This inequality can be restated as:

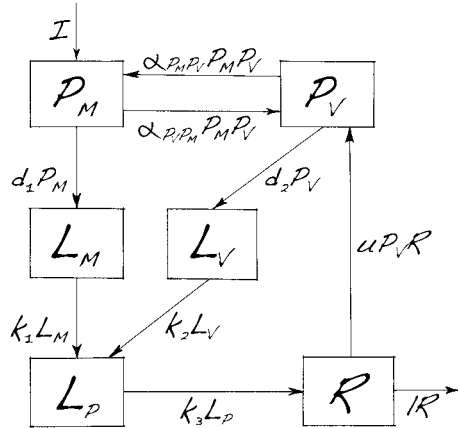
$$I + \delta \mu S \bar{\sigma}_i (\bar{R}_i^* - R_i^*) > R_i^*$$

or that $R_i^* < \text{inputs to } R_i^*$, which is analogous to what we have seen in the previous models. Loreau (1998) further shows that this inequality determines the maximum stable species richness imposed by the mass balance of nutrient flux through the system and the effect of each species on it through attributes which determine per capita flux rates. This is a very remarkable contribution to the current debate on biodiversity and ecosystem properties because it establishes rigorous and testable criteria by which we can determine how biodiversity, species life history traits, and nutrient input–output budgets determine the rate of nutrient cycling and by implication productivity of an ecosystem, a contribution which perhaps has not had the attention it deserves. Species richness (S) and the total nutrient cycling rate $\left(\sum_{i=1}^S u_i R_i P_i \text{ or } \sum_{i=1}^S k_i L_i \right)$ can both be predicted *a priori* from the constraints of mass balance and species life history traits.

Segregating uptake of nutrient inputs from uptake of the inorganic resource pool

Pastor et al. (2002) developed and analyzed a model of species coexistence and nutrient flux in peatlands in which mosses and vascular plants draw on different pathways and pools of nutrient fluxes (Fig. 12-8). Ombrotrophic peatlands, otherwise known as bogs, are portions of peatlands in which the accumulation of several meters of peat have raised the rooting zone above the regional water table (Wright et al. 1992). The only inputs of nutrients to bogs are in precipitation, which is effectively

Fig. 12-8 Nutrient cycling in a peatland model, after Pastor et al. (2002), for Eq. 12.12.



absorbed by mosses (P_M , Urban et al. 1988) and eventually transferred to peat through moss litter. In contrast to mosses, the growth of the vascular plant nutrient pool is determined by the vascular plant nutrient capital itself (P_V) with an intrinsic uptake rate u (uptake per unit R and per unit P_V) and by R . Live moss and vascular plant nutrient capitals are transferred to litter (L_M and L_V) at specific rates d_1 and d_2 , respectively. Litter decays at rates intrinsic to each species (k_3 and k_4 for moss and vascular plant litters, respectively) and eventually becomes peat (L_P). The peat in turn decays and releases nutrients into the inorganic nutrient pool at rate k_5 . Nutrients leave the system from the inorganic nutrient pool at a proportional rate l .

Because mosses and vascular plants do not draw nutrients from a common pool, they do not directly compete for nutrients (although mosses delay the availability of incoming nutrients to vascular plants). However, mosses and vascular plants reduce the growth of each other's nutrient capital through competition coefficients (α_{VM} = effect of mosses on vascular plants and α_{MV} = effect of vascular plants on mosses) representing competition for other resources, most likely competition for light through the shading of mosses by vascular plants or the engulfing of meristemic tissues of vascular plants by vertical growth of the moss carpet (Crum 1992).

The differential equations are:

$$\begin{cases} \frac{dP_M}{dt} = I - d_1 P_M - \alpha_{P_M P_V} P_M P_V \\ \frac{dP_V}{dt} = u P_V R - d_2 P_V - \alpha_{P_V P_M} P_M P_V \\ \frac{dL_M}{dt} = d_1 P_M - k_1 L_M \\ \frac{dL_V}{dt} = d_2 P_V - k_2 L_V \\ \frac{dL_P}{dt} = k_1 L_M + k_2 L_V - k_3 L_P \\ \frac{dR}{dt} = k_3 L_P - u P_V R - l R \end{cases} \quad (12.12)$$

There are two equilibria, a moss monoculture:

$$P_M^* = \frac{I}{d_1}, L_M^* = \frac{I}{k_1}, P_V^* = L_V^* = 0, L_P^* = \frac{I}{k_3}, R_M^* = \frac{I}{l} \quad (12.13a)$$

and an equilibria where mosses and vascular plants coexist:

$$\begin{aligned} P_M^* &= \frac{\alpha_{P_M P_V} l d_2 + \alpha_{P_V P_M} I u}{(\alpha_{P_M P_V} + \alpha_{P_V P_M}) d_1 u - \alpha_{P_M P_V} \alpha_{P_V P_M} l}, L_M^* = \frac{d_1 P_M^*}{k_1}, \\ P_V^* &= \frac{(d_1 u - \alpha_{P_V P_M} l) I - l d_1 d_2}{\alpha_{P_M P_V} l d_2 + \alpha_{P_V P_M} I u}, L_V^* = \frac{d_2 P_V^*}{k_2}, \\ L_P^* &= \frac{k_1 L_M^* + k_2 L_V^*}{k_3}, R_C^* = \frac{(\alpha_{P_M P_V} + \alpha_{P_V P_M}) d_1 d_2 + \alpha_{P_V P_M}^2 I}{(\alpha_{P_M P_V} + \alpha_{P_V P_M}) d_1 u - \alpha_{P_M P_V} \alpha_{P_V P_M} l} \end{aligned} \quad (12.13b)$$

The moss monoculture equilibrium (Eq. 12.13a) is very simple because the chain of fluxes from I to P_M to R is simply a chain of linear, donor-controlled flows with no feedbacks. The coexistence equilibrium (Eq. 12.13b) looks daunting, but notice that the equilibria of some compartments have complicated expressions which are shared by other compartments. For example, P_M^* , L_M^* , L_P^* , and R^* all contain the term $(\alpha_{P_M P_V} + \alpha_{P_V P_M}) d_1 u - \alpha_{P_M P_V} \alpha_{P_V P_M} l$ in their denominators thereby suggesting positive correlations between these state variables. Furthermore, the term $\alpha_{P_M P_V} l d_2 + \alpha_{P_V P_M} I u$ appears in the denominators of both P_V^* and L_V^* but in the numerators of both P_M^* and L_M^* . This suggests that the equilibrium nutrient capitals of vascular plant and vascular plant litter vary inversely with those of moss and moss litter.

The coexistence equilibrium value for each compartment also includes the competition coefficients between vascular plants and mosses. Therefore, competitive interactions between mosses and vascular plants for resources other than nutrients affect not only their relative nutrient capitals, but also the sizes of the peat and inorganic nutrient pools. This suggests a quantitative interaction between different resources, a phenomenon we will explore further in Chapter 14 on stoichiometry.

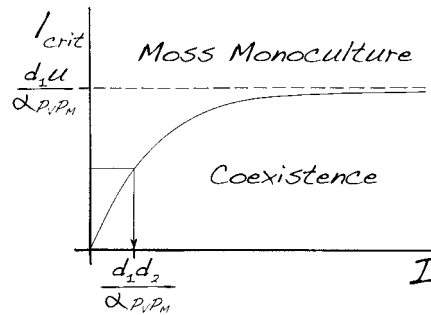
For the moss monoculture equilibrium, the eigenvalues of the Jacobian are $-l$, $-d_1$, $-k_1$, $-k_2$, $-k_3$, and $C/(d_1 l)$ where C is defined by:

$$C = (d_1 u - \alpha_{P_V P_M} l) I - l d_1 d_2 \quad (12.14)$$

Stability of the moss monoculture solution thus depends on whether $C < 0$ or $C > 0$. When $C < 0$, the moss monoculture is also globally stable since it is a purely linear system. C appears to be a function controlling the nature of a bifurcation between solutions. Stability of the moss monoculture thus depends on the parameters on the r.h.s. of Eq. 12.14.

Note that C is also the numerators of P_V^* and L_V^* for the coexistence equilibrium. If $C < 0$, then vascular plant coexistence equilibrium nutrient capital (P_V^*) is negative and it is therefore biologically unattainable but when $C > 0$ the moss monoculture is unstable and both P_V^* and P_M^* for the coexistence solution are positive. If $C = 0$, the two equilibria coincide. We therefore have a transcritical bifurcation that separates the two solutions, depending on this critical combination of parameters

Fig. 12-9 A two-parameter transcritical bifurcation between moss monoculture and moss–vascular plant coexistence in Eq. 12.12 as determined by the peatland’s input–output budget.



in C. Thus, the change in stability of the moss monoculture is accompanied by a “crossing” of the moss monoculture to the coexistence equilibrium. The existence of either community depends on the input–output budget of the system (I and l), on the values of species traits that control rates of transfers within the system (d_1 , d_2 , and u), and on the ability of mosses to engulf vascular plants and starve them for light ($\alpha_{p_v p_M}$).

Let’s solve for a critical export rate, l_{crit} at $C = 0$ in terms of the input rate (I) and all the other parameters. This will give us some idea of how the input–output budget of the system is controlled by the species’ traits. Surprisingly, l_{crit} is described by a Michaelis–Menten equation (Fig. 12-9) with the limit of l_{crit} equal to $d_1 u / \alpha_{p_v p_M}$ as I increases; the approach to the limit is determined by the half-saturation constant $d_1 d_2 / \alpha_{p_v p_M}$:

$$l_{crit} = \frac{\frac{d_1 u}{\alpha_{p_v p_M}} I}{\frac{d_1 d_2}{\alpha_{p_v p_M}} + I} \quad (12.15)$$

Therefore, l_{crit} describes a link between ecosystem input–output budgets (I and l), species’ traits (d_1 , d_2 , $\alpha_{p_v p_M}$, and u), and the stabilities of the two equilibrium communities. The bifurcation from a stable moss monoculture to a stable moss–vascular plant coexistence solution therefore arises from interactions between the input–output budget with the life history traits of the species in the communities. When $l > l_{crit}$, then moss monoculture is globally stable, but if $l < l_{crit}$, then coexistence is (locally) stable. This is biologically reasonable since when l is small, nutrients stay in the system and are available for vascular plant uptake.

This interaction of ecosystem nutrient budgets with life history characteristics has some interesting consequences. For example, it implies that fertilizing a bog (increasing I) should shift the ecosystem to the right of l_{crit} , thereby allowing vascular plants to invade and increase at the expense of mosses. Such shifts in species composition from moss to vascular plant dominance have been experimentally observed in fertilized bogs and heathlands (Aerts and Heil 1993), although these studies have not related them to the input–output budget and specific life history traits.

Summary: what have we learned?

Simple resource uptake models without litter return tend to either have unstable fixed point equilibria or else the conditions for stability of the fixed point are rather

delicately balanced. In addition, in food web or community models without litter return, increasing the number of links in the model also increases the chance for instability (May 1975b). But adding litter return globally stabilizes linear, donor-controlled models of food webs and communities as well as creating a large domain of stability around fixed point equilibria in nonlinear models. There are critical values for decay rates to stabilize these models, and in general slower decay rates increase the chances for stability.

We also examined a new measure of behavior of perturbed systems, their reactivity (Neubert and Caswell 1997). A system is reactive if perturbations are initially amplified, even though they are eventually damped back to zero as the system returns to its equilibrium. The reactivity of a system does not depend on the eigenvalues of the Jacobian, as matrices with the same set of eigenvalues can be either reactive or nonreactive. Instead, reactivity is defined by the eigenvalues of the Hermitian matrix of the Jacobian. Although low decay rates increase the chances for stability, high decay rates initially amplify perturbations and therefore make the system reactive. Thus, decay rates must be less than a certain critical value for the system to be stable, but greater than another for it to be reactive. Decay rates in systems with litter return therefore control both the long term stability as well as transient responses after perturbations.

Litter return itself does not allow stable coexistence between species at fixed point equilibria if both species take up nutrients via Lotka–Volterra interactions between the plant and the nutrient. This may come as a surprise to many ecosystem ecologists because it would seem that the return of resources in litter and their recycling through the resource pool give the other species a “second chance” at capturing the resource – two species should be able to share nutrients by recirculating them through the inorganic nutrient pool through return of their litter. However, species can coexist at stable fixed-point equilibria if the resource pools are at least partially segregated or if the species tap into different pathways of flux, such as the input flux vs. the decay flux to the resource pool.

Open questions and loose ends

Community ecologists have only recently begun to recognize litter decay and nutrient content as a “life history” trait of a species (e.g., Chapin et al. 1996, and also Wardle 2002 and references therein). By transplanting litter from different species or judiciously removing species from a plot, the effect of a species’ litter on nutrient cycling and coexistence of other species can be experimentally determined (e.g., Hobbie 1996). Similar approaches can be taken to test various predictions in this chapter, especially the critical decay rates for stability and reactivity. Are these critical decay rates correlated with other parameters of the species? If so, then there are tradeoffs between litter decay (which controls turnover of dead vegetation), uptake rates and death rates (which determines turnover of live vegetation) and the input–output fluxes of an ecosystem. It is well known that decay rates themselves are related to the chemistry of the litter (Berg and McClaugherty 2003). It therefore would be interesting to find out if the critical decay rates predicted for stability of various equilibria are similarly related to a plant species’ litter chemistry.

We have implicitly assumed that the plant pools themselves are homogeneous and that there is only one pathway of litter return. This has greatly oversimplified the situation and is perhaps only true for unicellular organisms such as phytoplankton or perhaps nonvascular plants such as mosses. All higher plants have at least three

different tissues – roots, stems, and leaves – each with different turnover rates and decay rates. An analysis of litter return and decay within the constraints and tradeoffs of allocation to different plant parts needed to acquire different resources is greatly needed. But the models which have been traditionally used to examine issues related to decay of different plant parts have typically been complicated simulation models such as CENTURY (Parton et al. 1988) and LINKAGES (Pastor and Post 1986). Because of their complexity, it is difficult to relate them to the results presented here on how litter affects ecosystem stability and reactivity. In order to do that, it is necessary to reduce them to simple differential equation form which can be linearized. A start on this approach has been made by Bolker et al. (1998), which could prove promising.

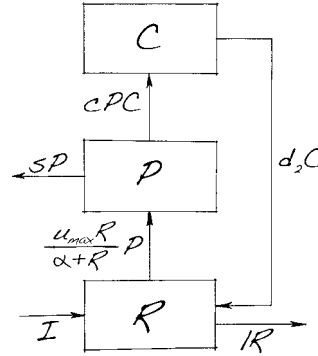
13 Consumer regulation of nutrient cycling

Almost from the beginning of life on earth, organisms have evolved which consume primary producers. Modeling consumer regulation of nutrient cycling raises two major problems. First consumers have a very different chemical composition than primary producers. Many plant tissues are carbon rich because they have lignified cell walls, but consumers, having cell membranes rather than thick cell walls, are less carbon and consequently more nutrient rich. This difference in chemistry imposes a strong limitation on the efficiency by which plant biomass can be converted to consumer biomass and is consequently a major factor determining the diet of consumers. We will take up the problem of modeling stoichiometry, or the ratios of nutrients and carbon in plants and consumers, in the next chapter.

The second problem that we face when incorporating consumers into our models is that some fraction of consumption as well as consumer bodies are returned to detritus as metabolic waste and as carcasses, respectively. Consumers therefore introduce an additional pathway of nutrient recycling to the inorganic nutrient resource pool, one which strongly affects system stability as well as nutrient availability and hence productivity of the primary producer (DeAngelis et al. 1989, DeAngelis 1992). How consumers affect the rate of nutrient cycling and the stability of ecosystems is the major problem that will concern us in this chapter.

In general, inclusion of nutrient recycling from consumers to the inorganic nutrient resource pool (via the intermediate step of detritus and decomposers that release the nutrients) increases the chances of stability in both closed (Austin and Cook 1974, Nisbet and Gurney 1976, Parker 1978) and open (O'Neill 1976) food web or ecosystem models, but much depends on the functions chosen to model nutrient uptake by plants and consumption of plants by consumers (DeAngelis et al. 1989). Empirical studies suggest that herbivores can either increase (McNaughton et al. 1997) or decrease (Pastor et al. 1993b, Ritchie et al. 1998) rates of nutrient cycling, so questions of when and how consumers affect rates of nutrient cycling are not easily settled one way or the other by experiments. In fact, it is more likely that under certain conditions consumers increase rates of nutrient cycling while under other conditions they decrease it, but finding those conditions empirically is a daunting task. As we shall see, mathematical analysis of nutrient fluxes can shed a great deal of light on what these conditions are. But without some guidance from mathematical models, one is forced to empirically search the number of possible configurations of herbivores, plants, and soil properties almost blindly. Because of the complications that consumers introduce into the stability and rates of nutrient cycles, there is perhaps no other problem in ecology which benefits more from mathematical investigations than consumer–plant–resource interactions in ecosystems.

Fig. 13.1 Nutrient cycling in a resource–plant–consumer system, modeled by Eq. 13.1.



One consumer–one plant species systems

As usual, we will begin by investigating the simplest possible system (Fig. 13-1), one with a single nutrient resource pool open to the environment (R), a single homogeneous plant compartment (P), and a single consumer (C). We will analyze a model of nutrient cycling in a phytoplankton–consumer pelagic system first constructed by O'Brien (1974) and analyzed further by DeAngelis (1992). First, like DeAngelis (1992), let's assume (since we have a model of a pelagic system) that dead plants which are not consumed sink to the sediments without decay at a per capita rate s and are buried (Fig. 13-1). Next, let's assume that all metabolic waste and carcasses of consumers (zooplankton) are immediately recycled back into the nutrient compartment, thus enabling us to do away with the detritus compartment. These two assumptions maximize the control consumers have on nutrient cycling rates. Relaxing these assumptions will only weaken the effect of the consumer, resulting in other solutions between this system and a system without a consumer.

The model is then:

$$\begin{cases} \frac{dR}{dt} = I + d_2C - \frac{U_{\max}R}{\alpha + R}P - IR \\ \frac{dP}{dt} = \left(\frac{U_{\max}R}{\alpha + R} - cC - s \right)P \\ \frac{dC}{dt} = (cP - d_2)C \end{cases} \quad (13.1)$$

The equilibrium solutions are:

$$R_0^* = \frac{I}{l}, \quad P_0^* = C_0^* = 0 \quad (13.2a)$$

$$R_1^* = \frac{s\alpha}{U_{\max} - s}, \quad P_1^* = \frac{I(U_{\max} - s) - ls\alpha}{s(U_{\max} - s)}, \quad C_1^* = 0 \quad (13.2b)$$

$$R_2^* = \frac{I - \frac{d_2s}{c}}{l}, \quad P_2^* = \frac{d_2}{c}, \quad C_2^* = \left[\frac{U_{\max}R_2^*}{\alpha + R_2^*} - s \right] / c \quad (13.2c)$$

(The subscripts indicate the number of trophic levels above R with nonzero values). The first equilibrium is a familiar one which we have seen in all other open systems with a nutrient resource: an “empty lake” consisting only of the nutrient resource without biota.

Let’s first examine the equilibrium with a plant but without a consumer (Eq. 13.2b). This will serve as a “null hypothesis” of a system without a consumer, so that we can then examine what happens when we add a consumer. Note that this solution is similar to one of the equilibria (Eq. 11.16b) of the Armstrong and McGehee (1980) model (Eq. 11.15), with d_1 in Eq. 11.15 and 11.16b replaced by s .

Although in Eq. 13.2a R^* is controlled entirely by its inputs and outputs, in Eq. 13.2b these no longer exert any control over R^* . Instead, R^* is controlled entirely by parameters pertaining to the plant (s , α , and $U_{\max}$). By contrast, the plant equilibrium is partially controlled by inputs and losses to R , namely by the parameters I and l .

To determine the conditions for the plant to invade the “empty lake” equilibrium, we can use the same technique that we used in Chapter 9 to determine when a competitor can invade a monoculture of another species when that species is at its equilibrium. That is, we assume that $P = \varepsilon$, $C = 0$, and $R = R_0^*$ and then ask under what conditions $dP/dt > 0$, which are:

$$\varepsilon \left(\frac{U_{\max} R_0^*}{\alpha + R_0^*} - s \right) > 0 \quad \text{or} \quad \varepsilon \left(\frac{U_{\max} \frac{I}{l}}{\alpha + \frac{I}{l}} - s \right) > 0$$

This inequality holds when the term in parentheses is positive, which we can solve for any parameter in terms of the others. Since the ecosystem is ultimately subsidized by the nutrient loading via the inputs, I , let’s solve this inequality for a minimum critical value of I that allows the plant to invade an empty lake (Eq. 13.2a). If this critical value of inputs allows a plant to successfully invade the empty lake, then we should obtain the second equilibrium with P_1^* . This minimum critical input is:

$$I > I_{\text{critical}, P_1^*} = \frac{ls\alpha}{(U_{\max} - s)}$$

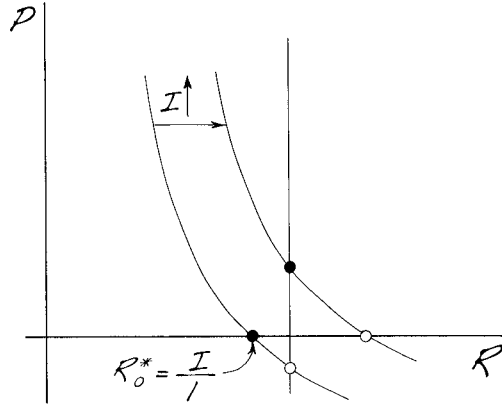
Once I matches this critical input, then Eq. 13.2a coalesces with Eq. 13.2b. To check this, insert $I_{\text{critical}, P_1^*}$ into R_0^* and you will see that we then recover the expression for R_1^* . So, once this critical value of I is exceeded, $P \rightarrow P_1^*$ and $R_0^* \rightarrow R_1^*$. We therefore seem to have a transcritical bifurcation between Eq. 13.2a and 13.2b at this critical value of I .

To check whether this is a transcritical bifurcation, let’s look at the nullclines for R and P when $C = 0$, which will show us how Eq. 13.2a and Eq. 13.2b relate to each other (Fig. 13-2). These nullclines are:

$$R \text{ nullcline: } P = \frac{\frac{I\alpha}{R} + I - l\alpha - lR}{U_{\max}}$$

$$P \text{ nullcline: } R = \frac{s\alpha}{U_{\max} - s}$$

Fig. 13-2 Nullclines, equilibria and their stability, and a transcritical bifurcation between Equilibrium 1 and 2 of Eq. 13.1.



The P nullcline is simply a vertical line, but the R nullcline is a monotonically decreasing function with increases in R . The R nullcline intersects the R axis (and $P = 0$) when $R = I/l$, or R_0^* . When I/l is to the left of the vertical P nullcline, then Eq. 13.2a is stable, P_1^* is negative, and Eq. 13.2b (R_1^*, P_1^*) is unstable. (Note the resemblance of these nullclines to those of Eq. 12.3 shown in Fig. 12-3 for plant uptake with litter return.) As I increases, the R nullcline moves to the right and the two equilibria approach each other and collide when

$$\frac{I}{l} = \frac{s\alpha}{U_{\max} - s} \quad \text{or} \quad I = \frac{ls\alpha}{U_{\max} - s} = I_{\text{critical}, P_1^*}$$

At this point, the two equilibria exchange stabilities: Eq. 13.2b becomes the stable equilibrium and Eq. 13.2a becomes unstable.

Thus, there is a critical value of input required before the plant can invade and maintain stable coexistence with R . Once the plant invades, however, R_1^* is completely determined by s , α , and $U_{\max}$, which are all parameters for the growth of P . The plant now exerts complete control “from the top down” on R .

Now let's turn our attention to Eq. 13.2c with the consumer coexisting with the plant and the resource. The equilibrium value for the plant is now entirely determined by the traits of the consumer (c and d_2), just as before in Eq. 13.2b the equilibrium value for the nutrient resource was entirely determined by the traits of the plant. Thus, the top-most trophic level exerts control over the equilibrium of the next lower trophic level, whether it is the plant over the nutrient resource (Eq. 13.2b) or the consumer over the plant (Eq. 13.2c).

Now let's ask under what conditions a consumer can invade Eq. 13.2b by assuming that $C = \varepsilon$ and $P = P_1^*$ and asking under what conditions $dC/dt > 0$:

$$\varepsilon(cP_1^* - d_2) > 0 \quad \text{or} \quad \varepsilon \left[c \frac{I(U_{\max} - s) - ls\alpha}{s(U_{\max} - s)} - d_2 \right] > 0$$

This inequality holds when the term in square brackets > 0 , which we can also solve for I , yielding:

$$I > I_{critical, C_2^*} = \underbrace{\frac{ls\alpha}{(U_{max} - s)}}_{I_{critical, P_1^*}} + \frac{sd_2}{c}$$

If we insert $I_{critical, C_2^*}$ into P_1^* , we at first get a very ugly expression, but with some patience and algebra almost everything magically cancels (try it!) and we are left with d_2/c , which is P_2^* . So, when input exceeds another, higher critical level required to support a consumer, then $P_1^* \rightarrow P_2^*$ and $C \rightarrow C_2^*$. We therefore have another trans-critical bifurcation between Eq. 13.2b and Eq. 13.2c at this higher critical value of I .

Exercise 13.1

Show that the coexistence equilibrium of R , P , and C is stable (all eigenvalues have negative real parts) when $I > I_{critical}$ (see Appendix for MatLab commands to do this).

Thus, as before where there was a critical input required before the plant can invade the abiotic nutrient resource, there is also a greater critical level of input required before the consumer can invade a plant–nutrient resource system, augmented above that required for the plant to invade the empty lake by the amount sd_2/c . This makes sense because now additional nutrient loading, proportional to the ratio of the per capita death (d_2) rate of the consumer to the per capita proportion of P consumed (c), will be needed to support both the plant and the consumer, compared with just the plant alone. In general, each new trophic level added to an ecosystem requires greater nutrient loading to the inorganic resource pool. We now have a mechanistic basis for the arguments of Oksanen and colleagues (Oksanen 1983, 1988, Oksanen et al. 1981) that the length of food chains depends on the productivity of the environment, only we can now state it in terms of the mass balance of nutrient inputs in relation to the life history traits of the different species rather than simply in terms of the producer's carrying capacity (K) as we saw in Chapter 8.

One can also calculate a minimum per capita proportion of P which must be consumed (c) for the consumer to invade a plant–resource system:

$$c > c_{critical, C_2^*} = \frac{d_2s(U_{max} - s)}{I(U_{max} - s) - ls\alpha}$$

Increasing c causes R_2^* to increase monotonically and asymptotically to I/l , which is R_0^* (Fig. 13-3). Therefore adding a consumer results in depressing R^* to levels below

Fig. 13-3 Effects of increasing consumption rate, c , on the size of R_2^* for Equilibrium 3 of Eq. 13.1.

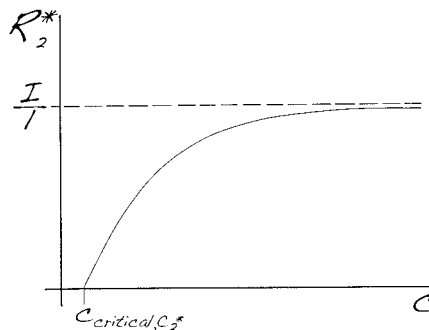
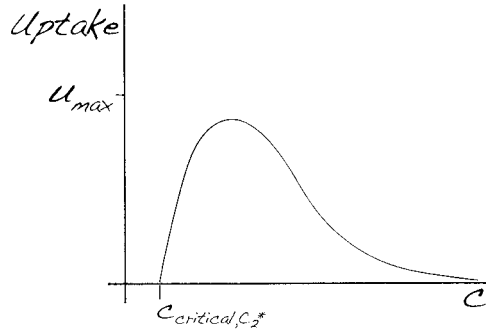


Fig. 13-4 Effects of increasing c on uptake rate at Equilibrium 3 of Eq. 13.1.



that of the system without a consumer (implying that more R now must be diverted not only to P but also to C), but increasing the rate of consumption causes R^* to rise (because at equilibrium a high consumption rate must be balanced by a high return rate of metabolic waste and carcasses from C to R).

How does increasing c control the flux rate through R , rather than simply the equilibrium pool size of R ? At equilibrium, the nutrient cycling rate equals the per capita uptake rate times P_2^* , or:

$$Uptake = \frac{U_{\max} R_2^*}{\alpha + R_2^*} P_2^*$$

Inserting the solutions for R_2^* and P_2^* , this reduces to:

$$Uptake = \frac{U_{\max} d_2 (Ic - d_2 s)}{(I\alpha + I)c^2 - d_2 s}$$

which is a concave downward parabola with uptake equal to 0 at $c_{\text{critical}, C_2^*}$ (Fig. 13-4). There is, therefore, both a minimum value of c required for stable coexistence of the consumer with the plants and the resource and a second value of c at which the rate of nutrient cycling through the system is maximized. The maximum value of uptake therefore exists at a unique value of $c > c_{\text{critical}, C_1^*}$.

This model therefore supports the logic behind the concept of “herbivore optimization” of nutrient cycles (McNaughton 1979), a conclusion that also arises from examination of more complex models as well (Loreau 1995, deMazancourt et al. 1998), although perhaps “maximization” is a better term than “optimization” since optimization usually refers to maximizing or minimizing a ratio rather than simply maximizing a rate or a pool size. In an analysis of a more complicated model, deMazancourt et al. (1998) show that consumers can maximize nutrient cycling rates subject to two conditions: (i) the proportion of nutrients exported from the system in herbivore biomass (through migration, for example) be less than the proportion lost from all other compartments combined, and (ii) inputs of nutrients to the system must exceed some threshold level in order to allow for flux to plants and then to the consumer. Eq. (13.1) meets condition (i) since we assumed that all consumer biomass and metabolic waste are recycled within the system; and it meets condition (ii) from the considerations of increased nutrient loading required for consumer invasion and stable coexistence. There is also an implicit assumption

not only in this model but also those of Loreau (1995) and de Mazancourt et al. (1998) that the plant pool is homogeneous and does not consist of different species with different uptake rates and consumer preferences, a point to which we shall return in the next section.

These results have important implications for how we can interpret different experiments on the effect of consumers on ecosystem nutrient cycles. Suppose, for example, we did some experiments and the results conform to the model results here, *but we didn't do the mathematics first to understand these complications*. Suppose we chose to measure R^* to characterize the nutrient cycle. If we did an experiment in which we excluded consumers and compared this to an identical situation with consumers, we would conclude that consumers significantly decrease R (Fig. 13-3). However, if we did an experiment in which we had a gradient of consumption rates from low to high (for example by judicious choice of which species of zooplankton you used) and regressed R against c , then we would conclude that increasing per capita consumption rates increases R (see also Fig. 13-3). Different experimental designs would lead to different conclusions and we would need both the enclosure and the gradient experiment to understand what was happening to R .

On the other hand, suppose we chose to measure the flux through R to plants to characterize the nutrient cycle. We would then do an experiment in which we again varied c as above and measured maximum plant uptake rate of R , perhaps through the use of an isotopic label for R . We would then conclude that increasing consumption first increases the rate of nutrient cycling to a maximum followed by a depression of nutrient cycling rate as consumption increased further (Fig. 13-4).

Failure to distinguish the effects of the absence/presence of a consumer from the effects of the intensity of consumption, and the effects of a consumer on the size of a nutrient pool vs the effects on the flux rate through that pool can be (and has been) the cause of much confusion about the effects of consumers on nutrient cycles. As we have seen, excluding consumers on R does not have the same effect as changing the intensity of consumption, and increasing consumption rate affects R differently than the flux through R . Working through the logic of our thinking by using even simple models such as this helps clarify the situation, or at least demonstrates that we must be clear and precise about what sort of experimental manipulation we are doing and what ecosystem property we are measuring.

These solutions also have considerable implication for the current debate over "top-down vs. bottom-up control" over ecosystem structure and functions. In every case, the traits of the top-most trophic level completely control the equilibrium of the next lower trophic level. But there is also a critical level of input to the "bottom-most" trophic level, namely R , required for entry and stable persistence of the top-most trophic level. Therefore, the top-most trophic level cannot exert any control unless this critical input to the bottom-most trophic level is exceeded, but once it invades, it completely controls the equilibrium of the next lower trophic level. In general, this reliance of the top-most trophic level on critical inputs to the bottom of the food web and in turn exerting complete control over the next lower trophic level continues with successive additions of trophic levels (DeAngelis 1992). Therefore, both bottom-up and top-down controls operate simultaneously, and each depends to some extent on the other.

The problem with the debate over “top-down or bottom-up” controls lies in the use of the conjunction *or*. Use of this conjunction (or its equivalent, *vs.*) poses a false dichotomy that is impossible to resolve experimentally because the system simply does not work that way. More precisely posed questions are: (i) how do the resource at the bottom-most level and the consumer at the top-most trophic level interact simultaneously to control ecosystem nutrient fluxes and standing stocks, within constraints of mass balance of inputs and outputs, and (ii) which traits do these controls depend upon? Failure to understand the interdependence of bottom-up and top-down controls and their simultaneous operations has been the cause of much confusion in the literature. Once again, an analysis of a simple mathematical statement of the problem helps clarify the situation in a way that no amount of verbal sparring or empirical tests in the absence of theory can do.

Exercise 13.2

Based on the patterns we have just seen about the input requirements for adding another trophic level and the control the highest trophic level has on all lower trophic levels, we can now make a conjecture about input requirements for adding a predator and how the predator controls lower trophic levels:

- There should be an even higher critical level of I that allows invasion of a predator.
- The predator should completely control the consumer equilibrium.
- This should release the plant equilibrium from complete control by the consumer.
- The plant should now be partially under the control of nutrient inputs.

Add a predator which preys on the consumer with a Lotka–Volterra (proportional harvesting) functional form to Eq. 13.1, find the coexistence equilibrium for all four compartments (see Appendix for MatLab commands to do this), and determine the critical input required to add a predator using the techniques shown above. Does the conjecture hold?

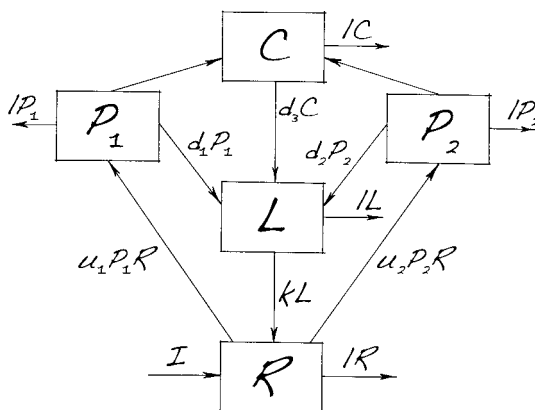
Exercise 13.3

Modify Eq. 13.1 to include plant litter return by diverting the proportion of P that sinks (sP) to the resource pool R . Find the equilibria, determine their stability, and find bifurcations that separate them.

Consumer regulation of plant species coexistence and nutrient cycles

The above model and much of the literature have assumed that the plant compartment is homogeneous with respect to uptake rates, consumer preference, and decay rates. But it is also readily apparent that herbivores do not forage through a puréed cream of vegetable soup, but instead forage in a cafeteria, combining or avoiding plant species in various proportions to meet their dietary needs (Stephens and Krebs 1987), often for a particular nutrient such as nitrogen which is in critical supply for protein metabolism (White 1993). Plant species themselves differ with respect to uptake rates and decay rates of their litter, and in fact, there are broad correlations between herbivore preferences, per capita uptake rates, and the chemical quality of plant tissues that control decay rates (Grime et al. 1996). This is because the same chemical factors that determine decay rates (such as lignin or nitrogen contents; see Berg and McClaugherty 2003 for comprehensive review of the literature) also determine digestion rates for the simple reason that digestion by many herbivores (especially mammalian ruminants) is microbially mediated. Thus, plant species with high per capita uptake rates have high nutrient concentrations. They are often preferred by herbivores because they are more nutritious but they also produce litter that decomposes quickly. By contrast, slowly growing plant species with low nutrient concentrations are often avoided by herbivores and also

Fig. 13-5 A model of nutrient cycling in which a consumer feeds in a cafeteria composed of two plant species for Eq. 13.3, after Pastor and Cohen (1997).



produce slowly decaying litter (Bryant and Kuropat 1980, Coley et al. 1985). The net effect can be a consumer-induced shift in plant dominance from species which cycle nutrients rapidly but are preferred to plant species that cycle nutrients slowly and are avoided, causing the rate of nutrient cycling to decrease (Pastor et al. 1993b, Ritchie et al. 1998).

But this poses obvious problems for the consumer. First, if the rate of recycling of nutrients declines, then so does the productivity of the preferred plant species. Second, the consumer cannot graze or browse the preferred plant species to the point where it is competitively excluded by the unpreferred species: the consumer must maintain a coexistence of both species. Coexistence can also benefit the consumer if the species which is not a preferred food source nevertheless provides some nonfood benefits, such as thermal cover or predatory cover.

Furthermore, herbivores must achieve all this within the constraints of the mass balance of input–output budgets of the ecosystems wherein they dwell. Thus, nutrient cycling considerations pose additional problems for the foraging animal beyond simply their immediate nutritional requirements traditionally considered by foraging theory.

Pastor and Cohen (1997) investigated these issues with a model of an inorganic nutrient pool (R), two plant species (P_1 and P_2), a consumer (C), and a litter/decomposer pool (L) (Fig. 13-5). Lotka–Volterra uptake functions are assumed for the plants (with per capita uptake rate u_i) and for consumption (with per capita consumption rate c_i) by the consumer. All detritus is returned to a single detritus pool but at different rates (d_i), these rates subsuming both death rate and decay rate. In order to reflect the empirical correlations between uptake rate, decay rate, and consumption rate mentioned above, assume that whichever plant species has the higher u_i , also has the higher c_i and d_i . To make things simple, we will also assume that per capita export rates from each compartment are equal, an assumption equivalent to saying that some portion l of the total nutrient capital [$l(R + P_1 + P_2 + C + L)$] is lost from the system without specifying a particular mechanism (leaching, transport by fluid medium, or migration) as the main loss mechanism. The mass balances of each compartment are given by:

$$\begin{cases}
 \frac{dR}{dt} = I + kL - (u_1P_1 + u_2P_2)R - lR \\
 \frac{dP_1}{dt} = u_1P_1R - c_1P_1C - d_1P_1 - lP_1 \\
 \frac{dP_2}{dt} = u_2P_2R - c_2P_2C - d_2P_2 - lP_2 \\
 \frac{dC}{dt} = c_1P_1C + c_2P_2C - d_3C - lC \\
 \frac{dL}{dt} = d_1P_1 + d_2P_2 + d_3C - kL - lL
 \end{cases} \quad (13.3)$$

There are (besides the usual “empty resource” or “empty lake” solution), one equilibrium solution for each plant species in monoculture without a consumer. As we have seen in the two previous chapters, two plant species cannot coexist on a single resource if they each have Lotka–Volterra uptake mechanisms. However, here there is an equilibrium solution in which the two plants can coexist with the consumer and the decomposer. Therefore, in the presence of a consumer, coexistence of two plant species with Lotka–Volterra uptake functions is possible, a solution also found by Grover (1997) in a model without litter return. The coexistence equilibrium expressions are all unwieldy except for that of the consumer compartment, which is:

$$C^* = \frac{u_1(d_2 + l) - u_2(d_1 + l)}{u_2c_1 - u_1c_2} \quad (13.4)$$

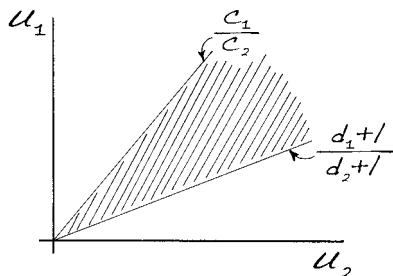
In order for C^* to be positive, the numerator and denominator must have the same signs. Therefore, for $C^* > 0$, the following conditions must be satisfied:

$$\frac{d_1 + l}{d_2 + l} < \frac{u_1}{u_2} < \frac{c_1}{c_2} \quad \text{or} \quad \frac{d_2 + l}{d_1 + l} < \frac{u_2}{u_1} < \frac{c_2}{c_1}$$

These ratios indicate the consumer’s preference for one plant over another (c_i/c_j), the relative difference in per capita uptake rates of the two plants (u_i/u_j), and the relative differences in death and decay rates (d_i/d_j). These conditions are satisfied whether: (i) P_1 is preferred by the consumer and cycles nutrients faster so long as $d_1 < u_1 < c_1$, or (ii) P_2 is preferred and cycles nutrients faster, again so long as $d_2 < u_2 < c_2$. This is a result of the symmetry of the differential equations for P_1 and P_2 .

Let’s therefore assume that P_1 is the preferred plant and cycles nutrients faster, simply because it is easiest to examine these constraints with the larger valued parameters for the preferred plant in the numerator. These inequalities imply that the consumer must select P_1 over P_2 to a greater extent than the plants differ in uptake rate, and the plant species in turn must be more different in uptake rate of the resource than in decay rate of their litter. Therefore, as you move up the food chain from plant uptake of a resource to consumer uptake of the plants, the next higher trophic level must be more discriminatory than the preceding one.

Fig. 13-6 The region of stable coexistence of the consumer with the two plant species (hatched area) for Eq. 13.3.



These inequalities determine a region in parameter space where coexistence is possible, which can be easily seen by multiplying through by u_2 , yielding

$$\frac{d_1 + l}{d_2 + l} u_2 < u_1 < \frac{c_1}{c_2} u_2$$

You can see that u_1 can be expressed in two ways as functions of u_2 , the decay ratios or preference ratios being the slopes, with c_1/c_2 being the greater of the two slopes. The consumer can coexist with the two plants only in the region between these two lines (Fig. 13-6). How can we broaden the possible region of parameter values that allows for coexistence? First, the consumer can be very discriminatory, increasing c_1 over c_2 , thereby steepening c_1/c_2 . Or the plants can differ less in decay rates, thereby making d_1/d_2 not as steep. Or export from the system, such as through emigration of the consumer, can increase because $\lim_{l \rightarrow \infty} (d_1 + l)/(d_2 + l) = 1$ so increasing l would decrease this ratio and broaden the region of allowable coexistence. However, exports are restricted by mass balance to meet the inequality $l(R^* + P_1^* + P_2^* + C^* + L^*) < I$, otherwise all compartments will go to zero. Therefore, a system with the widest allowable range of parameter values for coexistence is one with a consumer that exhibits a high degree of preference for the faster cyclor and with some (but not exceedingly high) leakage of the nutrient by migration. The importance of the migratory term for the consumer is surprising, and strongly suggests that to preserve coexistence of the two plant species, the consumer must move around in the landscape. We are thus on the edge of needing to introduce spatial dynamics into the system, which we will take up in Chapter 16.

Given the importance of the export term, it is reasonable to ask what happens if we close the systems, that is make both I and l equal to 0? To do this, we must go all the way back to the differential equations for the system (Eq. 13.3), remove the terms in I and l , and find new solutions for a closed system. For a closed system, C^* becomes:

$$\frac{d_1}{d_2} > \frac{u_1}{u_2} > \frac{c_1}{c_2}$$

These are exactly the opposite of the constraints for coexistence in the open system. In a closed system, in order to maintain coexistence the consumer must be less discriminatory between the two plants than the plants are in uptake rates, and the plants must differ even more in their decay rates than in the uptake rates. Without nutrient inputs, the constraint of conservation of the total nutrient capital of the system

reverses the conditions for coexistence. If the consumer is to remain stationary and not emigrate and if nutrient subsidies are negligible, then the consumer in a closed system cannot exhibit strong discriminatory preferences between plants. Thus, the pathways of nutrient cycling within a system and whether the system is open or closed determine the foraging behavior required of the consumer to maintain coexistence of the two plants.

The simplicity of this system matches that of the boreal forest, which basically consists of two functional groups of plants: fast cyclers which are generally deciduous plant species that are highly preferred by mammalian herbivores such as moose or snowshoe hare (Bryant and Kuropat 1980), have rapid uptake rates (Van Cleve et al. 1983), and rapid decay rates (Flanagan and Van Cleve 1983), and slow cyclers which are generally unpreferred conifers with slow uptake and decay rates (*ibid*). Using parameters for these species, Pastor and Cohen (1997) show numerically: (i) for stable coexistence of all compartments, the total rate of nutrient uptake at equilibrium must always be lower in the system with the herbivore than without, and (ii) increasing the herbivore's preference ratio depresses the rate of nutrient cycling.

Therefore, given a preference for the plant which cycles nutrients more rapidly, the net effect of the herbivore is to decrease rates of nutrient cycling by enabling the more slowly growing plant species with slower decay rates to coexist with the more rapidly growing plant species with faster decay rates.

However, there are ecosystems where herbivores do increase rates of nutrient cycling, most notably the Serengeti (McNaughton et al. 1997) and the American Great Plains (Hobbs et al. 1991, Frank and Evans 1997). It is significant, however, that the enormous herds of migrating mammalian herbivores in these grassland ecosystems consume almost all aboveground vegetation, thus effectively treating the plant compartment as a homogeneous pool. This conforms to the assumption of the first model in this chapter which, as we have seen, does allow for increased rates of nutrient cycling with increased consumption of plants (so long as the consumption and export in consumer biomass are not too high).

De Mazancourt and Loreau (2000) combine this model with the segregated nutrient pool model of Loreau (1998), which we examined in the previous chapter. When nutrient compartments for the two plant species were segregated then herbivores could maximize rates of nutrient cycling with increases in consumption rate up to a critical level of consumption, whereupon they declined, similar to our findings for the first model with homogeneous plant and resource pools. Therefore, herbivore maximization of nutrient cycling rates is possible even with a preference of the herbivore for the fast cycler so long as each plant species draws off its own segregated resource pool.

Finally, in a further analysis of the model considered here, Cohen et al. (2000) show that a mutant herbivore species can invade the system if it is evolved to prefer the slow cycler (such as spruce budworm in the boreal forest, for example). In that case, rates of nutrient cycling are increased and the highest rates of nutrient cycling are achieved at an evolutionarily stable equilibrium with maximum herbivore diversity.

The net result of these analyses and studies seems to be that consumer regulation of nutrient cycling, productivity, and coexistence of plants depends on whether the system is open or closed; the homogeneity of plant and resource compartments; and correlations amongst plant traits that determine uptake, consumptive preference, and

decay rates. In addition, effects on the size of R^* are not necessarily the same as effects on nutrient flux through R . Many effects of herbivores on nutrient cycling rates are possible and one must therefore be very careful in defining the system and determining which aspect of system behavior to focus on.

Summary: what have we learned?

Consumer regulation of nutrient cycling is not as simple as the oft-posed dichotomy of “top down vs. bottom up” controls. There is a critical level of input required before an additional trophic level can be added, so inputs to the “bottom” resource level control whether consumers can invade a system and persist. However, once the consumer invades, the trophic level below it is completely controlled by its parameters. Trophic levels “further down” are under less strong control of the consumer. Therefore, inputs to the “bottom” trophic level control the presence or absence of consumers, but consumers control the next lower level “from the top.”

How consumers affect nutrient availability depends on which measure of nutrient resource is considered. The size of the equilibrium nutrient pool increases monotonically with consumption rates, but the rate of nutrient flux through this pool increases to a maximum with increased consumption rates before declining with further increases. Care must be taken to distinguish between these two measures of nutrient resources when speaking of the effect of consumers on nutrient cycles.

Selective foraging by a consumer can establish stable coexistence between two plant species, but if the consumer prefers the species with greater uptake rates and decay rates (as is common for mammalian herbivores), then the rate of nutrient cycling is depressed as the less preferred species which cycles nutrients more slowly gains dominance. Moreover, the degree of preference between the two foods is partly determined by the input–output budget of the ecosystem: for an open system, a consumer can have greater preference for one species over the other than the two species have with respect to nutrient uptake, but in a closed system the preference differences must be less than the differences in uptake rates between the two plant species. Just as inputs and outputs determine whether consumers can invade a single plant system, so too do they determine the allowable limits of selectivity in multiple-plant species systems.

Open questions and loose ends

Most models of consumers and plants consider a single consumer compartment. These models need to be generalized to n -species models with many consumers feeding on many plant species, in other words to more complex food webs. Thébault and Loreau (2003) provide an excellent start to addressing this problem. They find that how a guild of consumers affects productivity and nutrient cycling depends on the structure of the food web. In general, they find that R^* declines with increased number of both plant and herbivore species while equilibrium plant biomass and hence nutrient uptake at equilibrium varies unimodally with increased number of plant and herbivore species, results qualitatively similar to those for the first model examined in this chapter. But if the number of plant species is held constant, then increasing herbivore species richness decreases R^* and equilibrium plant biomass, while herbivore biomass varies unimodally with herbivore species richness. The details of all cases also depend strongly on how the food web is connected. The richness of responses of their model may mirror the richness of responses seen in nature, and this paper is a promising start on generalizing consumer–plant–nutrient interactions to whole food webs.

In the models discussed in this chapter and most other models, the consumer is assumed to occupy a distinct trophic level and to be essentially a predator on the plant compartment(s) below it. But omnivores such as bears prey on both plants and herbivores and therefore occupy some sort of “intermediate” trophic level. Stable isotopes confirm that omnivores usually have signatures that are weighted averages of their foods, and thus occupy positions in food webs intermediate between organisms that feed on just one trophic level below them (Fry 2006). How do omnivores affect nutrient cycling rates? We saw in Chapter 10 that omnivory has a propensity to destabilize food webs, but those food web models did not have constraints of mass balance imposed on them. Does the mass balance of the flow of nutrients through foodwebs help stabilize them in the presence of omnivores? What sorts of bifurcations arise when omnivory is added to nutrient cycling models? As far as I am aware, these are completely open questions.

Humans have harvested both consumers and plants from consumer–plant–resource systems since before civilization. Indeed, we seem to have a particular fondness for herbivore flesh. While there has been much research on the effects of harvesting on the stability of predator–prey systems, no theoretical work has been done on the effects of harvesting on nutrient cycles in systems with consumers. It would be extremely interesting to take one of the models presented here or referenced in the literature and add harvesting terms such as in Exercise 8.3. Many game animals also live in systems in which the plants are also harvested, most especially in managed forests. How do harvesting of both game and the plants they subsist upon affect the cycling of nutrients? Are there critical ratios of the two harvest rates that maximize nutrient cycling rates?

Recently, there have been severe outbreaks of wildlife diseases, especially with the recognition of the role of prions in such epidemics (Grenfell et al. 2002). Models of the spread of diseases through populations have often been based on dividing a population between susceptible, infected, and recovered subpopulations. Such models are called SIR models, for susceptible, infected, and recovered (Grenfell et al. 2002). Because disease affects the metabolism of the herbivore, each of these subpopulations of the herbivore will have different consumption rates of plants. It would be very interesting to couple the SIR modeling techniques outlined in Grenfell et al. (2002) with one of the nutrient cycling models in this chapter to examine how diseases in wildlife populations affects the rates of nutrient cycling.

14 Stoichiometry and linked element cycles

Why stoichiometry is important and how it constrains allowable solutions of models

The stuff of which organisms are made (Wedin 1994) consists of abiotic materials which are combined only in certain ratios – indeed the ratios of different elements are almost an essential characteristic of each species (Sturner and Elser 2002). The characteristic elemental ratios of each species is called its stoichiometry. If we all were chemically identical (if we were all green slime) each species would have the same elemental ratios – the same stoichiometry – and feeding would simply be a matter of repacking the stuff from one form to another. But the stoichiometry of an organism is often very different from that of its food. This difference between the chemistry of a predator and its prey was implicit in the “metabolic conversion efficiency” term β of the Lotka–Volterra predator–prey equations: only a certain fraction, β , of the prey harvested by the predator is converted into predator biomass. Thus, a wolf with proportionally less skeletal tissue than a moose does not, over a period of time, consume the entire moose skeleton because it does not need all the calcium in the moose’s skeleton: although I have often seen moose bones with a few wolf tooth marks, the moose skeleton is almost always picked clean of the muscle tissue. Similarly, the moose requires proportionally less carbon than contained in the twigs that it browses upon, expelling the excess carbon in fecal pellets which are pressed sawdust almost devoid of any other nutrient other than carbon and the hydrogen and oxygen it is bonded to. Furthermore, organisms break down carbon compounds in their food through respiration to harvest the energy in their bonds and the carbon is expelled as CO_2 , so the ratio of carbon to that of the other nutrients, which are retained in proteins and other biomolecules, is necessarily altered.

This difference between the stoichiometry of an organism and its food means that some elements are present in the food resource in excess over the organism’s tissue requirements while other nutrients are deficient with respect to its tissue requirements. These deficient nutrients are likely to be limiting to the organism’s growth and development, including its reproductive potential.

Furthermore, as the organism grows, the allocation of growth to different tissues, each with a different stoichiometry, does not remain constant. For example, older and larger animals need proportionally more calcium-rich skeleton to support their mass than younger, smaller animals; as a tree grows, proportionally more of its biomass is in carbon-rich woody tissues compared to in the more nutrient-rich photosynthetically active leaves. Reproductive structures such as flowers are rich in phosphorus and antlers are rich in both phosphorus and calcium. Therefore, as the organism proceeds through the different phases of its life cycle, there are seasonal, annual, and decadal changes in its stoichiometry, and (in the case of antlers for most ungulates) even

sexual differences in stoichiometry. Different elements may be limiting at different life stages or for different sexes.

The organism overcomes the disparity between its stoichiometry and that of its food partly by expelling nutrients which are in excess and partly by being selective about what foods it should consume. The organism thereby alters the mass balance of the flux of each nutrient through its biomass in different ways: that is how it achieves a different stoichiometry than the food it consumes. Because certain nutrients which are deficient in the food are retained in an organism's biomass while other nutrients in excess in the food are expelled back to the environment, the organism also alters the stoichiometry of its environment. The magnitude and importance of this was first pointed out by Redfield (1958), who drew ecologists' attentions to the similarity of N:P ratios in marine phytoplankton ($N:P = 16$) and that of deep waters of the sea ($N:P = 15$), arguing that this was not an adaptation of the phytoplankton to their environment, but instead the result of phytoplankton importing atmospheric N (through N fixation to meet their demands) into an otherwise nitrogen-deficient ocean ecosystem.

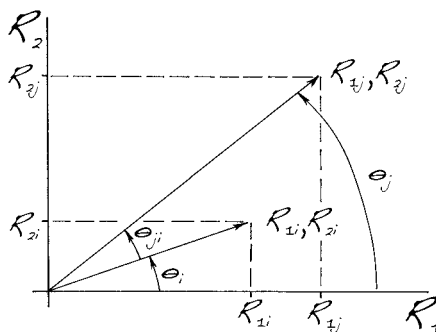
The stoichiometry of an organism therefore links multiple element cycles. As elements are transferred up the food web and then back to decomposers and the inorganic pool, there is a continual adjustment of these ratios through differential adjustment of diet, incorporation into tissues, and export back to the environment. It is hard enough to keep our thinking straight with single element cycles, but we can see that it is even more essential, but even harder, to consider linked element cycles.

Mathematical models of the stoichiometry of multiple element cycles are therefore essential to sharpen out thinking. Developing a consistent and comprehensive mathematical theory of the stoichiometry of linked element cycles is one of the cutting edge problems in mathematical ecology today, and the field is rapidly developing.

Elemental ratios and transformations in the phase plane

Consider a phase plane (Fig. 14-1) of two essential and limiting resources (usually nitrogen and phosphorus, but we will use R_1 and R_2 to keep things general). The axes represent masses of each resource in an organism's biomass or the biomass of the entire population. Therefore, each point in this phase plane represents a particular amount of R_1 and R_2 in a body or a population. The $R_1:R_2$ ratio is the slope of a line from the origin to that point. All other organisms that have the same $R_1:R_2$ ratio lie somewhere along this line. Thus, the phase plane consists of an infinite number of rays emerging from the origin, each of which define an $R_1:R_2$ ratio. The biomass of an individual (or population) determines its location on a particular ray. The

Fig. 14-1 The trigonometry of stoichiometry of two organisms.



location of a species in the $R_1 - R_2$ plane is therefore a vector with length $\sqrt{R_1^2 + R_2^2}$ at an angle θ with the abscissa. The stoichiometric ratio is the tangent of the species characteristic θ and the biomass determines the species characteristic vector length (Fig. 14-1).

Now consider two species: species j with (R_{1j}, R_{2j}) that consumes species i with (R_{1i}, R_{2i}) . The two species lie on different rays in the $R_1 - R_2$ plane (Fig. 14-1). It is readily apparent that the only way to alter the $R_1:R_2$ ratio from species i to species j is to alter the mass of either R_1 or R_2 or both – one cannot change a ratio without also changing amounts of at least one resource. Thus, changing the stoichiometric ratio from prey to predator involves changing the mass balance of at least one resource relative to the other. However, you can change the amounts of R_1 and R_2 in the organism without changing the ratio – simply multiply R_1 and R_2 by the same factor, thereby sliding it up and down the ray which defines a particular ratio.

There is therefore an asymmetry between ratios and mass balance: changing a ratio necessarily involves altering the amounts of at least one resource but it is possible to simultaneously change the amounts of the resources without changing the ratio. This asymmetry between ratios and amounts makes it exceedingly difficult if not impossible to separate the effect of resource ratios from resource amounts through a traditional ANOVA experimental design. Imagine doing an experiment in which you fertilized a plant with different amounts of R_1 and R_2 but in which you also wished to test the effect of different ratios of $R_1:R_2$ in the fertilizer on plant growth. Since altering the ratio of $R_1:R_2$ in the fertilizer cannot be achieved without changing the amount of at least one of the nutrients, the experimental manipulation of ratios is necessarily confounded with the manipulation of amounts. Different ratios cannot simply be treated as another “main effect” in an ANOVA experiment.

When the consumer (species j) eats the prey (species i) and converts the stuff of species i into its own stuff, it transforms the position of the point with coordinates (R_{1i}, R_{2i}) in the plane to the new point with coordinates (R_{1j}, R_{2j}) . This requires both a rotation and a translation in the $R_1 - R_2$ space. These two operations reflect different aspects of the biology of the two species.

To change the $R_1:R_2$ ratio from species i to species j , we need to rotate the vector defining species i from angle θ_i to θ_j , or through a change of angle $\theta_{ji} = \theta_j - \theta_i$. The ratio $R_{1i}:R_{2i}$ is the tangent of the angle θ_i and so $\theta_i = \tan^{-1}(R_{1i}/R_{2i})$. Therefore, the change of angle $\theta_{ji} = \tan^{-1}(R_{1j}/R_{2j}) - \tan^{-1}(R_{1i}/R_{2i})$. The angles θ_i and θ_j are characteristic of each species and angle θ_{ji} characterizes their stoichiometric interaction, or the stoichiometric imbalance between the two species due to their different tissue chemistries (Stern and Elser 2002, caption to their figure 8-1). There is a simple way to achieve this rotation and that is by multiplying the vector $[R_{2i}, R_{1i}]^T$ by a rotation matrix:

$$\begin{bmatrix} \cos \theta_{ji} & -\sin \theta_{ji} \\ \sin \theta_{ji} & \cos \theta_{ji} \end{bmatrix} \begin{bmatrix} R_{2i} \\ R_{1i} \end{bmatrix} = \begin{bmatrix} R_{2i,rotated} \\ R_{1i,rotated} \end{bmatrix} \quad (14.1)$$

Now we are on the ray defining the ratio $R_{1j}:R_{2j}$. These are the mass balance changes required to achieve a change in stoichiometry.

But we are not yet on the coordinate (R_{1j}, R_{2j}) of the consumer. To get to the consumer's coordinate we need to translate the rotated point up or down the ray

defining the ratio $R_{1j}:R_{2j}$ by a factor equivalent to the differences in individual or population biomass. This involves a second translational transformation defined by a diagonal matrix:

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \begin{bmatrix} R_{2i,rotated} \\ R_{1i,rotated} \end{bmatrix} = \begin{bmatrix} R_{2j} \\ R_{1j} \end{bmatrix} \quad (14.2)$$

where a is the ratio of the biomass of species j relative to species i , measured either as body sizes of individuals or biomass densities of the populations.

Therefore, the interactions between two species in multiple element space can be defined by two numbers: θ_{ji} , the stoichiometric interaction, and a , the ratio from prey to consumer biomass. Consuming prey and converting it into consumer biomass requires two different mass balance transformations: one to bring the chemistry of the prey in line with that of the consumer (the rotation) and the other to satisfy the biomass or body mass of the consumer (the translation). The stoichiometric angle of interaction can be calculated from the measured R_1 and R_2 contents through the use of $\tan^{-1}$ transformations as above. The two transformations can be combined as follows:

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \left(\begin{bmatrix} \cos \theta_{ji} & -\sin \theta_{ji} \\ \sin \theta_{ji} & \cos \theta_{ji} \end{bmatrix} \begin{bmatrix} R_{2i} \\ R_{1i} \end{bmatrix} \right) = \begin{bmatrix} R_{2j} \\ R_{1j} \end{bmatrix} \quad (14.3)$$

Changes in mass balance and stoichiometry between successive steps in a food chain can be phenomenologically modeled using Eq. 14.3 for each step from prey to consumer. This does not, of course, explain the mechanisms behind the mass balance and stoichiometry inherent in food webs. However, the rotational and translational matrices allow us to separate the changes forced by differences in stoichiometric ratios from those forced by differences in body size or biomass density at each step in the food web. These rotational and translation changes are precisely the phenomena that we wish dynamic models of linked element cycles to explain, and whether the models achieve the measured values of θ_{ji} and a can be a test of the model's accuracy.

Stoichiometry and competition between two species

Consider a single plant species with biomass P taking up two limiting nutrients, R_1 and R_2 as first proposed by Tilman (1980, 1982). The nutrients are supplied from large pools held in reserve in soil or sediment, S_1 and S_2 , respectively. The simplest assumption of the dynamics of the two nutrients in the absence of the plant is that they are proportional to the differences between R_i and S_i :

$$\frac{dR_i}{dt} = a(S_i - R_i) \quad (14.4)$$

Think of aS_i as a input rate of S_i to R_i (perhaps from decay of an organic matter pool S) and aR_i as a leaching loss rate. We are assuming for simplicity that the per capita decay and loss rates are equal, but strictly speaking the per capita decay and leaching loss rates, a , are equal only in a chemostat. In a chemostat, S is a large vat

of dissolved nutrient. The experimenter allows the nutrient to enter the chemostat at some fractional rate a by adjusting an input valve. Indeed, Tilman originally proposed this as a model of phytoplankton in a chemostat. We can rewrite Eq. 14.4 as $dR_i/dt = I_i - lR_i$ as in our previous models (where $I_i = aS_i$ and $l = a$), but as we shall see there is some advantage to writing it as we have because both R_i and S_i have the same units (mass per volume or mass per area) whereas I_i has the units mass per area or volume per time.

Now consider that each nutrient is taken up at a per capita rate c_i in proportion to plant biomass, so that the total uptake rate for each nutrient is:

$$u_i = c_i P \quad (14.5)$$

A consequence of this assumption is that the nutrients are obviously taken up in a constant stoichiometric ratio, namely $c_1:c_2$.

We can restate Eqs. 14.4 and 14.5 as vectors in $R_1 - R_2$ space:

$$\mathbf{I} = a \begin{bmatrix} S_1 - R_1 \\ S_2 - R_2 \end{bmatrix} \quad (14.6)$$

$$\mathbf{U} = P \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (14.7)$$

where $\mathbf{I}$ and $\mathbf{U}$ are the net input and uptake vectors in $R_1 - R_2$ phase space, respectively. At equilibrium, the $\mathbf{I}$ and $\mathbf{U}$ vectors must be of equal length and opposite sign, or

$$a \begin{bmatrix} S_1 - R_1^* \\ S_2 - R_2^* \end{bmatrix} - P^* \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (14.8)$$

By multiplying through terms, it is easy to see that the equilibrium point requires that:

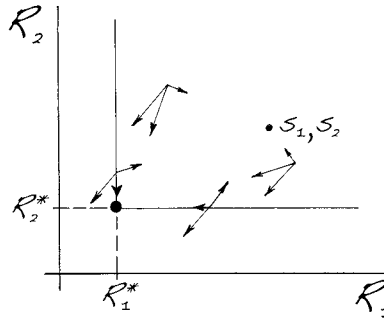
$$(S_1 - R_1^*) \frac{a}{c_1} = (S_2 - R_2^*) \frac{a}{c_2} \quad \text{or} \quad \frac{S_1 - R_1^*}{S_2 - R_2^*} = \frac{c_1}{c_2} \quad (14.9)$$

Thus, the “invasion” of an empty lake by a plant containing the two nutrients results in the equilibrium R_i^* being controlled by plant uptake rates (c_i), just as we have seen in the previous two chapters for a single nutrient.

We now plot the location of (S_1, S_2) in the $R_1 - R_2$ phase plane along with the equilibrium values and nullclines for R_1 and R_2 . Since the mass balance of each nutrient is independent of the other, the nullclines are mutually perpendicular in phase space, intersecting at the point (R_1^*, R_2^*) . If both nutrients also limit growth, then the equilibrium point is at the intersection of these two nullclines (Fig. 14-2).

The trajectory of (R_1, R_2) in this phase space is determined by the supply and uptake vectors at any point. The supply vector will always point to the coordinates of the supply pool (S_1, S_2) . Using the Pythagorean Theorem, the length of the supply

Fig. 14-2 Nullclines, the equilibrium point, and the vector field for two essential resources taken up independently.



vector is determined by the differences between S_i and R_i . The farther away (R_1, R_2) is from (S_1, S_2) , the longer the vector and the more the pull in that direction. The direction of the uptake vector is determined by the stoichiometric uptake ratio $c_1:c_2$, which is the slope of a vector from (R_1^*, R_2^*) at the corner of the two nullclines to the origin. Because the plant always takes up nutrients in the stoichiometric ratio $c_1:c_2$, the direction of the uptake vector for any point in phase space is parallel to the direction of the vector from (R_1^*, R_2^*) at the corner of the two nullclines to the origin. The length of the uptake vector at any point is determined by P at that point (Fig. 14-2).

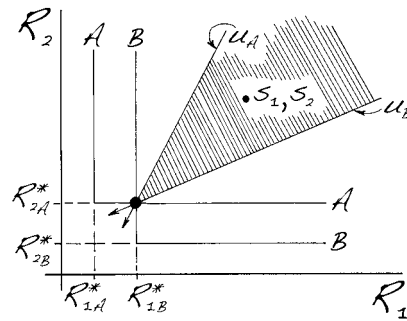
The plant equilibrium is stable only if $S_i > R_i^*$, and since the model is linear it is globally stable. This stability requirement specifies that the point (S_1, S_2) lies above and to the right of the two nullclines. If it lies below or to the left of the two nullclines, then $P^* = 0$ because $S - R^*$ will be negative for at least one of the nutrients (Fig. 14-2).

Because the stoichiometry of uptake is a fixed ratio, above and to the right of the nullcline the trajectories of (R_1, R_2) are straight lines towards one of the arms of the nullcline. The balance between the uptake and supply vectors moves the system towards the right-angled nullcline. Once the system lands on one of the arms of the nullcline, the uptake and input vectors combine to move it along the nullcline to the equilibrium point at (R_1^*, R_2^*) at the corner. Now, the line from the origin to the equilibrium has the slope R_1^*/R_2^* which is necessarily equal to $c_1:c_2$.

What happens if we have two species, A and B, with uptake rates for R_1 and R_2 of c_{1A} , c_{2A} , and c_{1B} , c_{2B} , respectively. Each species now has its own set of right-angled nullclines in $R_1 - R_2$ phase space. There are four possible outcomes, depending on where the nullclines fall with respect to one another and the relative direction of each species' uptake vectors. If the nullclines for one species fall completely above and to the right of the other ($R_{iA}^* < R_{iB}^*$ for both resources i or vice-versa), then the species with the lower equilibrium for both R_i always outcompetes the first and we have a monoculture at equilibrium. This simply extends the results we saw in Chapter 11 for two species competing for a single resource, in which the species with the lowest R^* outcompetes the other, to multiple resources.

But now suppose that $R_{1A}^* < R_{1B}^*$ and $R_{2B}^* < R_{2A}^*$. Then the nullclines cross and the point of crossing is the equilibrium point of the two-species system (Fig. 14-3). The supply vector still always points to the supply point (S_1, S_2) . But now, the uptake vectors for each species at this equilibrium are not parallel to their uptake vectors when they are in isolation. Instead the slopes of the uptake

Fig. 14-3 Stable coexistence of two species competing for two essential resources. The region of stable coexistence is hatched.



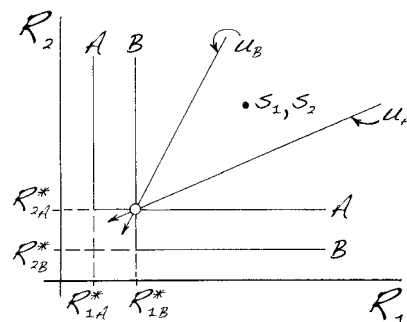
vectors are determined by the ratio of uptake of the two resources by both species simultaneously.

For each species, the resource most limiting to each species is the greater of R_1^* or R_2^* because it is reached first as the resources are drawn down from regions above and to the right of the nullclines. If the uptake vectors (u_A and u_B) for each species at the coexistence equilibrium are slanted towards the arm of the nullcline for the most limiting resource for that species, then each species preferentially draws down the resource most limiting to it and less limiting to the other.

If the supply point (S_1, S_2) lies between these two uptake vectors (u_A and u_B), as in the hatched region of Fig. 14-3, then the two species can coexist stably. By contrast, if the supply point lies outside this region but still above and to the right of at least one of the nullclines, then one species competitively displaces the other because one of the resources is drawn below the displaced species' corresponding R^* . And, as before, if the supply point lies below or to the left of both nullclines, then neither species can survive because $S_i - R_i^*$ will be negative for at least one resource for each species.

However, if the stoichiometric ratios of uptake cause the uptake vectors to slant towards the arm of the nullcline for the most limiting resource for the other species (Fig. 14-4), then each species preferentially draws down the resource most limiting to the other species. The coexistence equilibrium is then unstable. Which species then outcompetes the other depends on which one had greater initial biomass. Just as in the case of one configuration of the Lotka-Volterra competition model, which monoculture equilibrium is obtained depends on the initial ratios of population densities. We therefore have alternative stable equilibria, or bistability, as in the monoculture equilibria of the Lotka-Volterra competition models in Chapter 9.

Fig. 14-4 Unstable equilibrium of two species competing for two essential resources. The two species cannot coexist stably.



Therefore, not only do two resources allow coexistence of two species, the two species can widen their region of coexistence by having their stoichiometric ratios of uptake as different as possible, thus making the slopes of the uptake vectors as different as possible, and slanted towards the arm of their nullcline for the most limiting resource for that species. Different stoichiometries of uptake rate of two or more resources help avoid competitive displacement.

Exercise 14.1

How would you alter the differential equations for the resources in Eqs. 14.4 and 14.5 to reflect N fixation from the atmosphere for one nutrient and uptake from the water for the other nutrient, rather than uptake of both resources from the same aqueous pool?

Mechanisms of uptake of two resources

The previous model assumes that growth is a function of simple fixed stoichiometry and Liebig's Law of the minimum. Are there other mechanisms for multiple nutrient limitation of growth?

Recall from Chapter 11 that we derived a Michaelis–Menten model of nutrient limitation to productivity by introducing a delay in the arrival time, τ , of a nutrient to an uptake site in inverse proportion, c , to its concentration, R :

$$\tau = \frac{c}{R} \quad (14.10)$$

O'Neill et al. (1989) show that, using this technique, a wide variety of Michaelis–Menten functions can be developed for simultaneous effects of two nutrients on productivity. Let's begin by assuming that there are two nutrients, R_1 and R_2 , each with its own delay time, τ_1 and τ_2 , and proportionality constants, c_1 and c_2 . Assume that a reaction that produces biomass requiring both nutrients must wait a time τ_1 for the first nutrient to arrive and be taken up, then wait a time τ_2 for the reaction site to be ready for the next nutrient to arrive and be taken up. The reaction therefore proceeds in series. The total delay time is the sum of each individual nutrient's delay times, or:

$$\tau = \tau_1 + \tau_2 = \frac{c_1}{R_1} + \frac{c_2}{R_2} \quad (14.11)$$

In Chapter 11, we substituted Eq. 14.10 into

$$u = \frac{U_{\max}}{1 + U_{\max} \tau} \quad (14.12)$$

to yield the Michaelis–Menten function

$$\frac{dP}{dt} = \frac{P_{\max} R}{\alpha + R}$$

where $\alpha = cP_{\max}$ and dP/dt is the growth rate of plant biomass or nutrient pool, P . Similarly, if we substitute Eq. 14.11 into Eq. 14.12, after a bit of algebra to clean

things up (try it!) we arrive at the following Michaelis–Menten function for sequential uptake of two limiting nutrients and their incorporation into biomass as:

$$\frac{dP}{dt} = \frac{P_{\max} R_1 R_2}{\alpha_2 R_1 + \alpha_1 R_2 + R_1 R_2} \quad (14.13)$$

The half-saturation constants for each nutrient, α_i , are as before, but note now that their interaction results in the uptake and incorporation of each nutrient being determined by the half-saturation constant for the other nutrient as well as a cross product term $R_1 R_2$. O'Neill et al. (1989) call this the Additive model.

However, this model assumes that the order of arrival matters. If it doesn't then the arrival can be modeled as a random (stochastic) process. The simplest stochastic distribution for arrival times is a Poisson process. A Poisson process of arrival of two nutrients at random has three assumptions:

- 1 The numbers of changes from the arrival of one nutrient to the arrival of the other in successive nonoverlapping intervals are independent for all intervals.
- 2 The probability of exactly one change from the arrival of one nutrient to the arrival of the other in a sufficiently small interval $h = 1/n$ is $p = \nu h = \nu/n$, where ν is the probability of one change and n is the number of arrivals.
- 3 The probability of two or more changes in arrivals of two nutrients in a sufficiently small interval h is essentially 0.

In the limit of the number of arrivals becoming large, the resulting distribution is a Poisson distribution (the above conditions are adapted for two nutrients from the definition of Poisson processes in <http://mathworld.wolfram.com/Poisson-Process.html>).

Now assume that the two nutrients each have some probability density function $g_i(t)$ for their arrival times. Then the cumulative probability $G_i(t)$ for the arrival of each nutrient by time t is:

$$G_i(t) = \int_0^t g_i(t') dt'$$

Furthermore, the mean delay in arrival time for each nutrient is:

$$\tau_i = \int_0^{\infty} g_i(t') t' dt'$$

If this is the case, then $g_i(t)$ is:

$$g_i(t) = \frac{R_i}{c_i} e^{-\frac{R_i}{c_i} t} \quad (14.14)$$

As $t \rightarrow \infty$ you can easily see that $\lim g_i(t) = R_i/c_i$ and consequently the mean delay time in the arrival is the inverse or

$$\tau_i = \frac{c_i}{R_i} \quad (14.15)$$

Now one nutrient can arrive first at some time t and be in a “holding situation” while waiting for the other to arrive at some later time t' , where $t' > t$. Then the total delay time for both nutrients to arrive is:

$$\tau = \int_0^{\infty} g_1(t) dt \int_t^{\infty} g_2(t') t' dt' + \int_0^{\infty} g_2(t) dt \int_t^{\infty} g_1(t') t' dt' \quad (14.16)$$

The first term on the r.h.s. is the average over all arrival times when R_1 arrives before R_2 and the second term is the average over all arrival times for R_2 arriving before R_1 . Given Eq. 14.14 for each nutrient, then this can be solved for τ to be:

$$\tau = \frac{(R_1/c_1)^2 + (R_1/c_1)(R_2/c_2) + (R_2/c_2)^2}{(R_1/c_1)(R_2/c_2)(R_1/c_1 + R_2/c_2)}$$

By substituting this expression for τ into the Michaelis–Menten Eq. 14.12, we get (after cleaning up some messy algebra) the Poisson Arrival Time model:

$$\frac{dP}{dt} = \frac{P_{\max} R_1 R_2 (\alpha_2 R_1 + \alpha_1 R_2)}{(\alpha_2 R_1)^2 + (\alpha_1 R_2)^2 + R_1 R_2 (\alpha_2 R_1 + \alpha_1 R_2) + \alpha_1 \alpha_2 R_1 R_2} \quad (14.17)$$

You can see after some inspection that this is almost the same as multiplying the numerator and denominator of Eq. 14.13, the Additive model (where the sequence of arrival mattered), by $(\alpha_2 R_1 + \alpha_1 R_2)$ except that that would result in $2\alpha_1 \alpha_2 R_1 R_2$ in the denominator of Eq. 14.17 instead of just a single such term. The reason why there is only one $\alpha_1 \alpha_2 R_1 R_2$ in the denominator of Eq. 14.17 is because it doesn't matter which nutrient arrives first, whereas in the Additive model it did – having two such terms in the denominator of Eq. 14.17 would be double counting the arrive of a unit of R_1 and R_2 required to make a unit of biomass. Therefore, the two models are algebraically closely related, which is satisfying. The higher order terms in the Poisson Arrival Time model (Eq. 14.17) therefore represent “corrections” to the Additive model (Eq. 14.13) required by relaxing the assumption that one nutrient always arrives first and instead assuming that the order of arrival does not matter.

O'Neill et al. (1989) use this same delay time technique to derive a variety of other models as well, then tested the ability of each model to fit data on responses of *Monochrysis* algae to additions of phosphorus and vitamin B₁₂ (Droop 1974), *Stylosanthes humilis* (a legume) to additions of nitrogen and phosphorus (Gates and Wilson 1974), alfalfa and corn to additions of nitrogen and phosphorus (Heady et al. 1955), and various wetland graminoid species to additions of nitrogen and phosphorus (Shaver and Melillo 1984). In all cases, the Additive and Poisson Arrival Time models gave better fits to experimental data than all the other models, including Liebig's Law of the Minimum. Although the Additive and Poisson Arrival Time models could not be distinguished statistically, O'Neill et al. (1989) recommend using the Additive model because of its simplicity and because it approximates quite well the kinetics of the more theoretically sound but more complicated Poisson Arrival Time model.

The exercise of model fitting also resulted in some recommendations for the design of experiments. The estimation of parameter values and goodness of fit of each model were done through nonlinear regression, in which the sums of squares of residuals

are minimized not through exact formulas as in linear regression but through an iterative procedure starting with reasonable approximations of parameter values (see Seber and Wild 2003 for an introduction to these techniques). O'Neill et al. (1989) found that for all these datasets, it was most important to have reasonable estimates of $P_{\max}$ (maximum production or uptake rate) from the experimental data. This requires high enough levels of both nutrients so that further increases in production could no longer be observed – in fact, these high levels of nutrient additions are the more important controls to determine how nutrients limit growth because they are the levels when the two nutrients are not limiting. This is the true “control” for determining nutrient limitations. The more commonly used “control” of no additions of nutrients was the least important point of the dataset for estimating the parameter values. Secondly, suboptimal levels of both nutrients need to be represented in the dataset to obtain good estimates of the half-saturation constants. Indeed, the reason why these particular studies were chosen to test the models is because their datasets satisfied both these criteria, whereas a large literature search through many hundreds of nutrient addition experiments turned up numerous papers with datasets that did not satisfy these criteria and so could not be used to distinguish the different mechanisms of multiple nutrient limitation represented by the different models.

Adding stoichiometry to predator–prey models

A predator must sometimes contend with prey which have a different stoichiometry than it needs to build its own biomass. This is particularly so when the predator is an herbivore that consumes primary producers: plants have C:N or C:P ratios that can be an order of magnitude greater than that of the herbivore which consumes them, simply because plants often have lignified support tissues that are high in carbon and low in nutrients, whereas much of a consumer's body mass is dominated by soft tissues of high nutrient and low carbon concentrations. Even algae and the zooplankton that consume them have a wide disparity in C:nutrient ratios, although algae do not have lignified supports tissues. Thus, consumers face a major problem in converting their food, which is deficient in most nutrients relative to carbon, into its own biomass.

Loladze et al. (2000) approached this stoichiometric problem facing consumers by extending the Rosenzweig–MacArthur model of predator–prey interactions to include a limiting nutrient in addition to the carbon that makes up 50% of the producer and consumer dry masses. I will rewrite here the Rosenzweig–MacArthur model (Eq. 8.14) to make it easier to see how Loladze et al. (2000) modify it:

$$\begin{cases} \frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{K} \right) - f(N_1)N_2 \\ \frac{dN_2}{dt} = \beta f(N_1)N_2 - mN_2 \end{cases} \quad (14.18)$$

where

$$f(N_1) = \frac{cN_1}{d + N_1}$$

or the Michaelis–Menten/Holling Type II consumption function and N_1 is the biomass density (in carbon) of the producer and N_2 is the biomass density (also in carbon) of the consumer. K is therefore the equilibrium carbon in the biomass of the producer.

Loladze et al. (2000) modify these equations in the following ways: (i) they assume that the system, while open to carbon, is closed with respect to a limiting nutrient, nominally phosphorus, and that all of the phosphorus in the system (P) is distributed between the producer and the consumer; (ii) they add a second “carrying capacity” to the producer equation which depends on the amount of phosphorus available to the prey, which is the difference between the total phosphorus in the system and the phosphorus in the consumer; (iii) they modify the efficiency of converting producer biomass to consumer biomass by the degree to which the P:C ratio of the producer is less than that of the consumer (just as the P:C ratio of woody plant tissue that a moose browses is less than that of moose flesh).

If the system is closed to phosphorus flux, then total amount of phosphorus in the system, P , is fixed and

$$P \geq q_{N_2}N_2 + q_{N_1}N_1 \quad (14.19)$$

where q_{N_2} is the P:C ratio of the consumer and q_{N_1} is the minimum P:C ratio required by the producer (analogous to Droop’s minimum cell quota). If the total phosphorus in the system is distributed between producer and consumer, then $P - q_{N_2}N_2$ is the amount of phosphorus in the producer. Therefore, by subtraction in Eq. 14.19:

$$P - q_{N_2}N_2 \geq q_{N_1}N_1$$

or

$$1 \leq \frac{P - q_{N_2}N_2}{q_{N_1}N_1}$$

Now, the modified model becomes

$$\begin{cases} \frac{dN_1}{dt} = rN_1 \left(1 - \frac{N_1}{\min(K, (P - q_{N_2}N_2)/q_{N_1})} \right) - f(N_1)N_2 \\ \frac{dN_2}{dt} = \beta_{\max} \min \left(1, \frac{(P - q_{N_2}N_2)/N_1}{q_{N_2}} \right) f(N_1)N_2 - mN_2 \end{cases} \quad (14.20)$$

In the producer equation (dN_1/dt), the term $P - q_{N_2}N_2/q_{N_1}$ is the amount of phosphorus in the producer divided by the minimum phosphorus concentration required by the consumer. This is equivalent to the maximum possible carbon content of the producer given the amount of phosphorus available to it. This is the second and new carrying capacity, the other being of course our old friend, K , which is the maximum carbon content under density-dependent limitation of carbon flux due to the balance of photosynthesis and respiration without limitation by phosphorus. The $\min(\dots)$ function just implements the Liebig Law of the Minimum by choosing which

carrying capacity (phosphorus- or carbon-limited carrying capacity) determines producer growth in the absence of the consumer: the one with the least value.

In the consumer equation (dN_2/dt), the conversion efficiency, β , is also modified by the $\min(\dots)$ function that follows. The term $[(P - q_{N_2}N_2)/N_1]/q_{N_2}$ is the P:C ratio of the producer divided by the P:C ratio of the consumer (this is a ratio of a ratio, which gets very tricky, so you have to watch it like a hawk). If the P:C ratio of the producer is less than that of the consumer, then the consumer sees the producer as deficient in phosphorus relative to its needs and $[(P - q_{N_2}N_2)/N_1]/q_{N_2} < 1$. This term then reduces β from $\beta_{\max}$. When the P:C of the producer equals that of the consumer, then the consumer can convert the producer into its own flesh without phosphorus deficiency. Then $\beta = \beta_{\max}$ and we recover the original Rosenzweig–MacArthur predator equation.

The conservation of P in this closed system (Eq. 14.19) results in a new, second nullcline for the consumer:

$$N_2 = \frac{P}{q_{N_2}} - \frac{q_{N_1}}{q_{N_2}} N_1 \quad (14.21)$$

This nullcline intercepts the N_2 -axis at P/q_{N_2} and similarly intersects the N_1 -axis at P/q_{N_1} (Fig. 14-5). The slope of the nullcline is the ratio of the P:C ratio of the consumer to the minimum of the producer. This nullcline and the constraint of conservation of P in a closed system make the model a zero-sum game once P becomes limiting. In a zero-sum game, any gain by one of the players (the producer or consumer) comes at the expense of the other.

Suppose K is very small. Carbon is therefore limiting the producer and K is chosen by the $\min(\dots)$ function to limit producer growth. If K is small enough, then the consumer cannot survive. As K increases, the same thing happens as with the Rosenzweig–MacArthur model: we first pass through a transcritical bifurcation that allows coexistence between producer and consumer in a stable spiral node, followed by a Hopf bifurcation and the birth of a stable limit cycle with further increases in K .

As K increases even further, then the carbon in the producer also increases. Increasing K pulls the producer nullcline to the right, and the amplitude of the stable limit cycle also increases. When the producer nullcline, consumer P nullcline, and the limit cycle are all tangent to one another, the system passes through a saddle-node bifurcation (Fig. 14-5). This saddle node is actually a special type of saddle node, called a saddle node on an invariant circle (Fig. 14-6), or “snic” bifurcation (a stable limit cycle is a type of invariant circle, invariant because the system stays on it or returns to it after perturbation).^{*} The limit cycle becomes pinched between the parabolic producer nullcline and the P-limited consumer nullcline as this snic bifurcation is approached. Once the nullclines pass through the snic bifurcation, the limit cycle disappears.

With further increases in K , the snic bifurcation splits the equilibrium into a stable and unstable pair of coexistence equilibria. The stable equilibrium is a fixed-point spiral node (Fig. 14-5) and the stable limit cycle disappears. As K continues to increase,

^{*}Thanks to Harlan Stech, Department of Mathematics and Statistics, University of Minnesota Duluth, for pointing this out.

Fig. 14-5 Nullclines and bifurcations as K increases and P becomes limiting in the Loladze et al. (2000) model, Eq. 14.20.

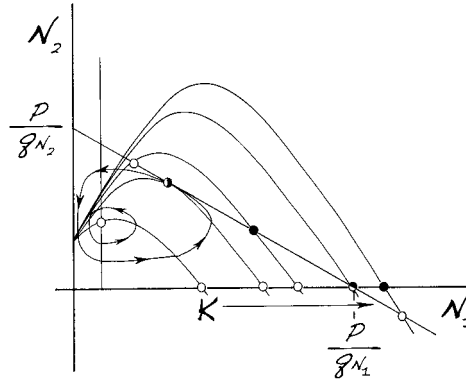
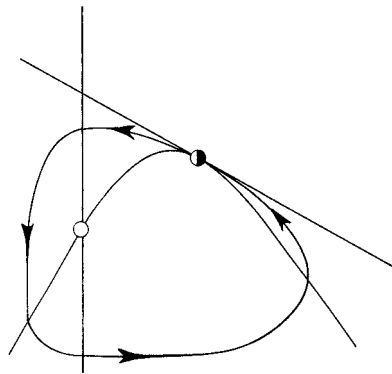


Fig. 14-6 A saddle-node on an invariant circle (“snic”) bifurcation in the Loladze et al. (2000) model.



the stable equilibrium slides downward toward K , where they eventually collide and exchange stabilities in a second transcritical bifurcation when $K = P/q_{N_1}$ (Fig. 14-5). Beyond this bifurcation, the coexistence equilibrium is in the fourth quadrant and is unstable; the consumer is now extinct because of low food quality (high C:P in the producer). The entire sequence of bifurcations as K increases is summarized in Fig. 14-7.

Thus, adding nutrient limitations to the Rosenzweig–MacArthur predator–prey model through stoichiometry and making the system closed to the fluxes of the nutrient (but not carbon) results in new bifurcations and extinction of the consumer either at low K (not enough carbon fixed by the producer to support higher trophic

Fig. 14-7 Bifurcation diagram for the Loladze et al. (2000) model as K increases.

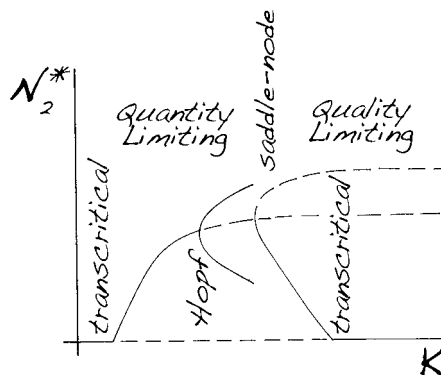


Fig. 14-8 Bifurcations in the Loladze et al. (2000) model caused by decreasing the total amount of P in the system.

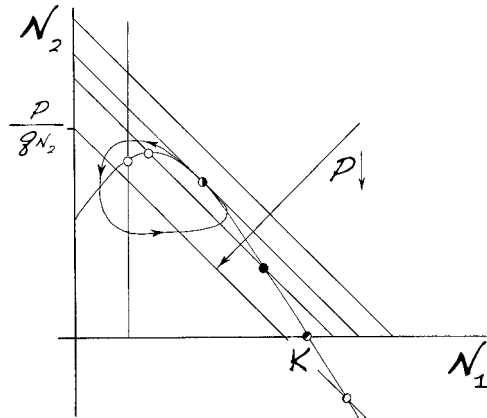
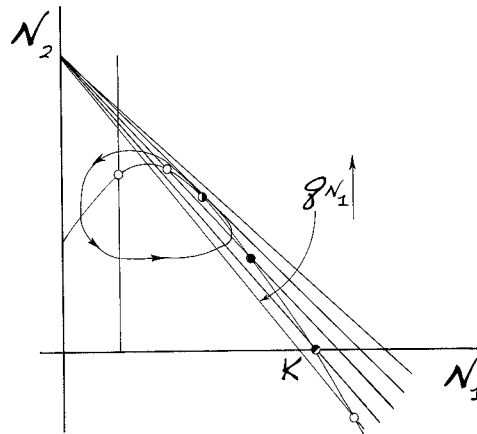


Fig. 14-9 Bifurcations in the Loladze et al. (2000) model caused by increasing q_{N_1} . Note that this rotates the P-nullcline for the consumer, as opposed to a translation caused by decreasing total system P as in Fig. 14-8.



levels) or at high K and low P (high C:P) in the producer and therefore poor quality food for the consumer.

These bifurcations can also be produced by reducing the amount of P in the system, which translates the second N_2 nullcline (Eq. 14.21) downward (Fig. 14-8). Alternatively, by increasing q_{N_1} (the minimum P:C ratio in the producer), the growth of the producer declines because of P limitation. Increasing q_{N_1} rotates the N_2 nullcline (Fig. 14-9). Thus, as in our analysis of stoichiometric interactions in the two nutrient phase plane with which we began this chapter, changes in the size of the nutrient pool cause translations (Fig. 14-8) and changes in stoichiometry cause rotations (Fig. 14-9) of the nullclines.

Effect of stoichiometry on producer-microbial interactions

Plant litter has a C:N ratio of 50–100, but microbes which convert this litter into their biomass have a C:N ratio of 5–10, an order of magnitude narrower. Therefore, for every mole of C in plant litter, there is about 10 $\times$ less N than the microbes need to build biomass. Microbes must therefore narrow the C:N ratio of plant litter to make their own biomass.

How can microbes do this? In two ways. First, distill off the excess carbon as CO_2 . In Chapter 3, we saw that this can be described by varieties of exponential decay

equations. In addition, microbes can import nitrogen from the surrounding environment, a process known as nitrogen immobilization (this also often occurs for phosphorus as well – see Berg and McClaugherty 2003). What does this do to the carbon and nitrogen mass balance of the decaying litter?

Aber and Melillo (1982) noticed empirically that the N:C ratio of litter increases linearly from its initial ratio ($N_0:C_0$) as the litter decays. This can be expressed as:

$$\frac{N}{C} = \frac{N_0}{C_0} + a \left(1 - \frac{C}{C_0} \right) \quad (14.22)$$

where $(1 - C/C_0)$ is the cumulative carbon or mass loss to time t . But what is a ? The units of a have to be N immobilized per g C respired or lost, and so a is a microbial nitrogen use efficiency relative to carbon respired.

Bosatta and Staaf (1982) take this a step further and derive a differential equation for the nitrogen in the decaying litter–microbial biomass complex. First, they insert the exponential decay equation for C ,

$$C = C_0 e^{-kt}$$

into Eq. 14.22, yielding (after a bit of cleaning up)

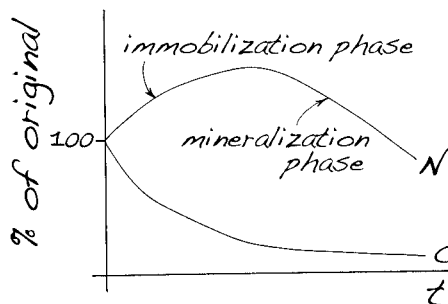
$$N = N_0 e^{-kt} + a C_0 e^{-kt} - a C_0 e^{-2kt} \quad (14.23)$$

Before we go further, let's think about what each of the three terms on the r.h.s. mean. The first term says that the nitrogen originally in the plant litter is released exponentially at the same rate as the carbon, perhaps by the breaking of C–N bonds in the plant material. The second term, which is positive, means that nitrogen is imported into the litter in proportion to carbon respired, the proportionality constant being equal to the nitrogen use efficiency of the microbes. But there is a third term that is negative, meaning that some of the imported nitrogen is also lost. This loss of nitrogen is from waste metabolism of nitrogen in the microbes which must be excreted as ammonium, otherwise ammonium can build up in the microbial cell to toxic levels. All three terms determine whether the net amount of nitrogen in the litter–microbe complex is increasing ($dN/dt > 0$) or decreasing ($dN/dt < 0$).

Now let's take the derivative of this equation to find the conditions under which there is net import or export of nitrogen from the litter–microbe complex. Recall from Chapter 2 that if $u = f(t)$, then $d(e^u)/dt = e^u du/dt$ and note that $u = -kt$. Then,

$$\begin{aligned} \frac{dN}{dt} &= -kN_0 e^{-kt} - a k C_0 e^{-kt} + 2a k C_0 e^{-2kt} \\ \frac{dN}{dt} &= [-kN_0 e^{-kt} - a k C_0 e^{-kt} + a k C_0 e^{-2kt}] + a k C_0 e^{-2kt} \\ \frac{dN}{dt} &= -k[N_0 e^{-kt} + a C_0 e^{-kt} - a C_0 e^{-2kt}] + a k e^{-kt} C_0 e^{-2kt} \end{aligned}$$

Fig. 14-10 Mass balances of carbon and nitrogen in decaying litter.



But the term in square brackets equals N from Eq. 14.23 and $C_0 e^{-kt} = C$, so by substitution

$$\frac{dN}{dt} = -kN + ake^{-kt}C \quad (14.24)$$

which is our final equation for nitrogen dynamics in decaying litter. The stoichiometric link between the carbon and nitrogen fluxes is provided by the parameter a , the nitrogen use efficiency of the microbes per unit carbon respired. The first term gives the rate of nitrogen outputs and the second term gives the rate of nitrogen imports due to immobilization. The nitrogen balance is due to the relative magnitudes of these two terms, the second of which depends on carbon and the microbial nitrogen use efficiency.

Now, what is the sign of dN/dt and how does it depend on a and t ? If $dN/dt > 0$ at $t = t_0$, then nitrogen will be initially immobilized and there will be a net increase in nitrogen in the litter even as carbon is respired (Fig. 14-10). This will happen when

$$0 < -kN_0 + akC_0$$

or when

$$\frac{N_0}{C_0} < a$$

Therefore, if $N_0/C_0 < a$, the amount of nitrogen in the litter-microbe complex will increase because N_0/C_0 is not sufficient to support microbial nitrogen demands per unit carbon respired and the microbes have to import additional N from the surrounding environment (Fig. 14-10). Measured rates of a range from 0.02 to 0.06. Given measured initial N/C ratio of 0.01 or less in many litters, nitrogen is immobilized during the early stages of decay of most litters (Berg and McClaugherty 2003). However, as $t \rightarrow \infty$, the second term on the r.h.s. of Eq. 14.24 $\rightarrow 0$ and net release of nitrogen, or net nitrogen mineralization, begins (Fig. 14-10). This is the point when the nitrogen in the litter-microbe complex becomes available to plants. Similar immobilization-mineralization dynamics also happen for phosphorus during decay of some litters (Berg and McClaugherty 2003).

These coupled nitrogen-carbon dynamics of decaying litter are not simply of interest to microbial ecologists but affect ecosystem dynamics as well. The period of net immobilization can be quite long – often several years – for many litters, especially for litters of initially low N:C ratios. Since microbes have many orders of magnitude greater surface area than plant roots and since uptake depends first on surface area, microbes essentially get whatever nitrogen they need first and plants get what is left over. Therefore, net initial immobilization results in a long delay in the release of nitrogen and hence a period of nitrogen limitation, often for several years after a cohort of litter falls. In Chapter 6 we saw how delays in differential equations can lead to oscillations if the delay is longer than $1/|\lambda|$, the characteristic return time of a system after perturbations. Pastor and Durkee Walker (2006) show that such delays can cause oscillations in a plant–litter–sediment model if a is large enough.

In Chapter 11, we saw that by widening its C:nutrient ratios (high nutrient use efficiency) a plant can maintain high biomass production per unit nutrient taken up. But models of Loladze et al. (2000) show this high C:nutrient ratio also means poor food quality for herbivores. High nutrient use efficiency therefore simultaneously decreases losses of the plant to consumers as it increases biomass production. But the plant eventually runs into self-imposed nutrient limitations because the high C: nutrient ratios of the litter imposes severe stoichiometric constraints on microbes decomposing the litter and recycling the nutrients back to inorganic form for the plants (the Bosatta and Staaf 1982 model). Conservation of matter and mass balance means that there is no free lunch: the increased growth and protection from consumption the plant gains by increasing its C:nutrient ratio is at least partially lost because the slow decay of its litter decreases nutrient availability.

Summary: what have we learned?

Organisms create their biomass with characteristic elemental ratios. Because these ratios differ for different organisms, the growth of an organism is partly determined by the stoichiometry of its food, and the stoichiometry of the organism also partly determines the stoichiometry of its environment. Consuming prey and converting it into consumer biomass requires two different mass balance transformations: one to bring the chemistry of the prey in line with that of the consumer (a rotation in the phase space of the two elemental resources) and the other to satisfy the biomass or body mass of the consumer (a translation in the phase space of the elemental resources). Adding stoichiometric constraints to plant–nutrient uptake models or to predator–prey models results in a rich variety of behaviors not seen in single element models, including coexistence on two resources and additional bifurcations and dynamics.

Open questions and loose ends

The mathematics of stoichiometry in ecosystems is one of the cutting edges of mathematical ecology, and the reader is strongly encouraged to read Sterner and Elser (2002) for many other experimental and modeling results and open questions. Here are some open questions that interest me in particular:

Open vs. closed systems – The Loladze et al. (2000) model is open to carbon flux but closed to phosphorus flux. Would the consumer go extinct at high amounts of food if the system were also open to phosphorus inputs? How does stoichiometry affect the behavior of open compared to closed systems? In Chapter 13, we saw that in a single nutrient model, a consumer must forage differently in an open system compared with a closed system if coexistence with two plant species is to be stable.

What happens when we add two nutrients with characteristic stoichiometric ratios for both the plant and the consumer? In an open system, does it matter which compartment the elements enter into and which they leave from? Two nutrients can enter a compartment in the ecosystem from the outside environment in any ratio, but when they leave in exported biomass (by emigration or harvesting, for example), they leave in the characteristic ratio for that compartment. Does it matter whether they are exported from the plant compartment or from higher trophic levels?

The min (. . .) function – Strictly speaking, Liebig's Law of the minimum only applies at an instant of time. When growth is integrated over a longer period, several resources may limit growth in sequence. In addition, the min (. . .) function produces a corner in the nullclines and in the differential equation. At this corner, the derivative is undefined. This is mathematically not very satisfactory. What happens when we replace the min (. . .) function with a smoother function, such as the multiple nutrient Michaelis–Menten functions derived by O'Neill et al. (1989)? Are there other, smooth functions of multiple nutrient limitations?

The allometry of stoichiometry – As most higher organisms grow, their stoichiometry changes because of increased requirements for support tissues to bear their weight. Such support tissues include wood for plants and skeletons for animals. Wood and skeleton each have very different stoichiometries than metabolic tissues: wood is much higher in carbon and skeleton is much higher in calcium and phosphorus than metabolic tissues. How does stoichiometry change with body size? What do these changes do to the stability of ecosystems? Is the distribution of body sizes in populations determined somehow by the mass balances of nutrient fluxes and the changing stoichiometries as body size increases?

Stoichiometric constraints throughout a complete nutrient cycle – Since two or more nutrients cycle simultaneously in ecosystems, maintaining mass balances throughout the nutrient cycles from plants, to consumers, to decomposers poses severe constraints on coexistence and productivity. For example, a plant can increase its productivity and decrease consumer biomass (and avoid being eaten) by increasing its C:nutrient ratios (which are also nutrient use efficiencies). But this then results in decreasing nutrient availability because the microbes which must decompose the plant litter to release the nutrients for plant uptake are starved for the nutrient relative to their stoichiometric requirements. This decreases microbial growth rate and increases microbial nutrient uptake rate, thus decreasing the rate of nutrient supply back to the plants and hence productivity. All stoichiometric models thus far have imposed stoichiometric constraints on only one or two trophic levels, not on the complete loop from plants to consumer to decomposers to the inorganic nutrient pools in soils. A model in which all compartments have stoichiometric constraints would be extremely interesting. I suspect that such a model would indicate that there are very tight constraints imposed on ecosystem structure and functions because of simultaneous mass balances. It would be extremely interesting if the relative distribution of equilibrium biomass and nutrients among different trophic levels arose largely from these constraints.

Part 4

Populations and ecosystems in space and time

15 Transitions between populations and states in landscapes

Throughout this book, we have assumed homogeneity of the environment and distribution of organisms. But the spatial environment is not homogeneous: landscapes are dynamic because processes and states of populations and ecosystems are not fully synchronized and because processes spread over them and change the distributions of populations and ecosystems. Such processes include fires and other disturbances, the flow of water downhill, and the wanderings of animals in search of food, amongst many others.

The theory of landscape dynamics is a rapidly developing field. It is not possible to cover the full range of problems of mathematical approaches to landscape dynamics in this and the next chapter, and you are referred to Tilman and Kareiva (1997) and Dieckmann et al. (2000), which are collections of very good papers on these problems. Instead, what I wish to do is to show you how the techniques we have used to analyze the dynamics of populations and ecosystems – such as matrix models, eigenvalues and eigenvectors, coupled differential equation models, and bifurcations – can also be used to gain some understanding of the dynamics of populations and ecosystem processes in space as well as time.

This chapter will focus on matrix models of landscapes and models of metapopulations. In these models the dynamics of the different states of the population or ecosystem across the landscape are modeled with respect to one another although the precise spatial location of a population or ecosystem is not explicitly considered. In the next chapter, we will explicitly consider the spread of populations across the landscape by means of diffusion equations, which will be derived explicitly from some basic assumptions of how organisms move across landscapes.

Markov Chains of landscapes

To model a landscape, we first need a mathematical object that can stand as an operational definition or metaphor for what you can see out of the window of an airplane, for example. Think of (or define) a landscape as a bounded surface composed of a set of n elements distributed among m classes (one of which can be an “empty” class with no living things) that change from one class to another during some time interval τ with certain probabilities, p_{ij} . The elements can be squares on a checkerboard, in which case they are called pixels, and the classes could be different vegetation types or land uses. The conversions of one class to others on such a landscape can be modeled as a matrix of transition probabilities:

From:

$$\text{To: } \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix} \quad (15.1)$$

where each p_{ji} denotes the probabilities of transition from class i to class j across m classes. This is known as a Markov matrix, and the system of equations in which this matrix multiplies a vector of states (here, acreages) of different classes is known as a Markov chain:

$$\mathbf{m}_{t+\tau} = \mathbf{A}\mathbf{m}_t \quad (15.2)$$

All columns in a Markov matrix must add to 1 to conserve probabilities, and hence to conserve acreages: either a given element remains in the same class during time τ with probability p_{ii} or it changes to one of the other specified classes with probability p_{iji} . All elements in the landscape belong to one of the classes. A first-order Markov process is any process in which the current state of the system depends only on the state in the previous time step.

Transition probabilities can be calculated in a number of ways. Horn (1975) constructed transition probabilities of species in a hardwood forest by enumerating the species of an overstory tree and the species of an understory tree judged most likely to replace the overstory tree when it died. A pleasant walk through the woods with a notebook is all that Horn needed to gather data. However, with the advent of geographic information systems a map from one date (derived, for example, from remotely sensed images, e.g., Hall et al. 1991, Wolter and White 2002) can be superimposed onto another from a different date, and the number of pixels that have changed over some time interval τ from i to j can be enumerated and transition matrices thereby parameterized. For a map with a large number of pixels (N), the maximum likelihood probability for a given transition is then:

$$p_{ij\tau} = \frac{n_{ij\tau}}{\sum_{i=1}^n n_{ij\tau}} \quad (15.3)$$

In other words, if you constructed a matrix of numbers of pixels that changed from i to j , in the limit as N is large the probabilities can be calculated by dividing each entry by the sum of the column in which it resides. Although the transition probabilities are derived from maps of changes from one class to another across a large number of locations, the transitions themselves are the average over all locations where a given transition has taken place (for example, from birch to spruce forest). Put another way, although the data used to parameterize the model is spatially explicit, the transition probabilities are averaged over all locations in which they occur. They do not depend on what is happening in the adjacent pixels.

Exercise 15.1

What would a Markov matrix look like if no transitions to different classes took place – that is, if every location stayed in its same class during time step τ ? What is the more general term for this matrix?

What if τ is not some convenient time interval, such as often happens when using historical data such as air photos, inventories, etc? Then the probabilities have to be normalized to some convenient time interval, such as years, decades, or centuries. This is not done by dividing the probabilities by τ and multiplying by the length of the time interval because the columns would no longer add to 1 and acreage would no longer be conserved.

The way to think about this is to imagine each transition as being a decay of one class into another. If the probabilities are constant, then the decay rates are constant. We now have to set up the exponential decay equations for the proportion of the landscape in each class $i(P_i)$ and solve them for different time steps. Here's how:

$$\frac{dP_i}{dt} = \sum_{j=1, j \neq i}^m k_{ij} P_i$$

where P_i is the proportion of the landscape in i and k_{ij} is the instantaneous rate of change (i.e., “decay”) from i to j . Now the proportion of i remaining that did not transfer to j after time interval τ is:

$$P_{i\tau} = e^{k_{ij}\tau}$$

and the proportion of i that “decayed” (transferred) to j in τ is:

$$P_{ij\tau} = 1 - e^{k_{ij}\tau}$$

Rearranging:

$$e^{k_{ij}\tau} = 1 - P_{ij\tau}$$

Now take logarithms of both sides and solve for k_{ij} :

$$k_{ij} = \frac{\ln(1 - P_{ij\tau})}{\tau}$$

Now substitute back into the equation for the proportion of i that decayed to j and solve for the desired time interval t :

$$P_{ijt, i \neq j} = 1 - e^{[\ln(1 - P_{ij\tau})/\tau]t}$$

$$P_{iit} = 1 - \sum_{i \neq j, i=1}^m P_{jit} \quad (15.4a)$$

or

$$\ln(1 - P_{ijt})/t = \ln(1 - P_{ij\tau})/\tau \quad (15.4b)$$

where P_{ijt} is the proportion (probability) of class i that decayed to class j in time interval t when τ is expressed in units of t and $P_{ij\tau}$ is the proportion remaining in class i . Note the symmetry of this equation.

Because the columns of a Markov matrix sum to 1, the dominant eigenvalue equals 1. By inspection of Eqs. 15.1 and 15.2, it can be seen that the normalized eigenvector associated with this eigenvalue will be the stable distribution of classes in the landscape, just as the dominant eigenvector of a Leslie matrix was the stable distribution of age classes in a population. Alternatively, you can think of each of the elements of the normalized dominant eigenvector as the probability that any pixel of land will eventually end up in a given state in the limit as $t \rightarrow \infty$.

Making this interpretation of the dominant eigenvector requires that the matrix has several properties that are subsumed under a general theorem called the Perron–Frobenius Theorem (Hohn 1973, Lancaster and Tismenetsky 1985):

- *Primitivity*, or the property that successive powerings of the matrix are positive. Since all probabilities are positive, this holds for all the matrices we will be considering.
- *Irreducibility*, or all states or classes must be connected to all others by a chain of $p_{ji} > 0$. By contrast, a reducible matrix has transient states that cannot be reached from any others.

If a matrix satisfies these properties then successive powering of the matrix times any vector will converge on the dominant eigenvector, or stable distribution of classes. Think of this as the climax distribution of classes in the landscape assuming the transition probabilities remain constant.

It is obvious that the matrices we will be using for landscape models are primitive, but whether or not a matrix is irreducible is not always so obvious. For example, in the following matrix:

$$\begin{bmatrix} p_{11} & p_{12} & 0 & 0 \\ p_{21} & p_{22} & 0 & 0 \\ 0 & 0 & p_{33} & p_{34} \\ 0 & 0 & p_{43} & p_{44} \end{bmatrix}$$

classes 1 and 2 can make transitions to each other but not to classes 3 and 4. Likewise, classes 3 and 4 can make transitions to each other but not to classes 1 and 2. The matrix is therefore reducible to two independent and irreducible matrices:

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \text{ and } \begin{bmatrix} p_{33} & p_{34} \\ p_{43} & p_{44} \end{bmatrix}$$

and the dominant eigenvectors of each are the stable distribution of the respective classes. It is as if we are dealing with two independent landscapes.

By contrast, in the following matrix:

$$\begin{bmatrix} p_{11} & 0 & 0 \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

class 1 is known as a *transient* class because it decays to 2 and 3 but cannot be reached from them. The matrix is therefore also reducible, but the submatrix of transitions between classes 2 and 3 is irreducible. If the dominant eigenvalue of the entire matrix equals that of this irreducible submatrix, then the dominant eigenvector of the entire matrix will still be the stable distribution of classes 1–3. Depending on how a matrix is assembled, it may not always be so easy to tell whether it is reducible or not by inspection.

Finally, if a class can be reached from all others but the probability of remaining in that class is 1, then the landscape eventually will converge on that class. That class is called an *absorbing* class.

As with Leslie matrices (Chapter 4), the ratios of the other eigenvalues to the dominant eigenvalue in Eq. 4.5 determine the transient behaviors and rates of approach of the population vector to the stable age distribution. The damping ratio ρ (the magnitude of the ratio $|\lambda_2|/\lambda_1$; Caswell 2000) determines the time required for the contribution of λ_1 to become x times as great as that of $|\lambda_2|$:

$$\rho^t = \left(\frac{\lambda_2}{\lambda_1} \right)^{t_x} = \frac{1}{x}$$

And by taking the logarithm of both sides and solving for t_x we get:

$$t_x = -\frac{\ln x}{\ln \rho}$$

The percent of convergence to the stable age distribution is $100 - (100/x)$. For 95% convergence to the stable age distribution, t is solved for $x = 20$ (i.e., $100 - (100/20) = 95$) at a specified ρ . This is equivalent to the contribution of $\lambda_1 > 20 \lambda_2$.

Exercise 15.2

Prove that the dominant (largest) eigenvalue of a general, irreducible 2×2 Markov matrix equals 1, hence the dominant eigenvector is the stable distribution of age classes.

Exercise 15.3

A landscape on the forest–prairie boundary has the following decadal transition probabilities (assume any woody community, including shrubs, is a “forest”):

	From:	
	<i>F</i> <i>P</i>	
To:	<i>F</i> $\begin{bmatrix} 0.75 & 0.1 \end{bmatrix}$	
	<i>P</i> $\begin{bmatrix} 0.25 & 0.9 \end{bmatrix}$	

Find:

- 1 λ_1 and λ_2 .
- 2 The damping ratio and the time (in decades) required for 95% convergence to steady state distribution of forest and prairie.
- 3 The steady state distribution of forest and prairie.

Exercise 15.4

A foraging snail eats algae on a rock. Assume the snail does not affect the subsequent growth rate of algae. In each day, the snail can convert 10% of the algal covering to bare rock, and in the same length of time the algae can grow back and cover 20% of bare patches.

Find:

- 1 The steady state distribution of bare rock and algae.
- 2 The time required to reach 95% of steady state.
- 3 If the density of algae is 1 mg/cm², the total surface area of the rock is 10 cm², and the snail requires 4 mg of algae per day, can the snail persist on this rock given the steady state amount of algae or does it have to move to a different rock?

Ecologists have criticized first-order Markov models because the transition probabilities for a given stage are not usually constant (see review by Baker 1989). But there can often be interesting reasons for the changes in transition probabilities. For example, beaver affect the transition probabilities between wetland types within their ponds as they build dams and occupy the flooded ponds behind them (Pastor et al. 1993a). If they are killed or abandon the pond, the dams, which require frequent maintenance, break and the pond drains to be replaced by a variety of seasonally flood or moist wetland meadows. Pastor et al. (1993a) calculated transition matrices for succeeding decades during the recovery of a beaver population in northern Minnesota. Transition matrices for valley bottoms (including the classes: not pond; flooded; seasonally flooded; and moist meadow) changed during the four decades from 1940 to 1980, as did their dominant eigenvectors. But in all transition matrices, the class “not ponded” was transient because all land in valley bottoms eventually becomes occupied by an active or abandoned beaver pond. This appears to have happened in the half-century it has taken for the beaver population to recover: by the 1990s, very little land in valley bottoms was not impounded or otherwise affected by beavers (Johnston and Naiman 1990). The series of transition matrices has a corresponding series of dominant eigenvalues. In this series of eigenvectors, as the stable proportion in flooded ponds increases, that in moist meadows decreases, and vice versa. The dynamics of the landscape are therefore dominated by increased occupation of valley bottoms by beaver ponds and meadows, and an alternation between flooded ponds and moist meadows as beaver ponds are occupied and abandoned (Pastor et al. 1993a).

Semi-Markov and age-structured models

Hobbs and Legg (1983) and van Hulst (1979) have suggested that transition probabilities within a stage depend on the age structure of the stage. To account for nonstationary behavior of transition probabilities, Acevedo et al. (1995, 1996) suggest replacing the Markov matrix A with another matrix Q such that

$$q_{ij} = p_{ij}F_{ij}(t)$$

where $F_{ij}(t)$ are functions describing sojourn or holding times in state i before transition to state j with probability p_{ij} . Usually some probability distribution, such as Poisson or gamma functions, are assumed for $F_{ij}(t)$. This is known as a semi-Markov process. These holding time functions essentially emulate “hidden” or implicit intermediate states between two or more states whose transitions are described by the first-order Markov model (Acevedo 1981). For example, in a

first-order Markov model, aspen forests can form a class. But aspen forests can in turn be classified as sapling, pole-sized, and timber stands, which represent increasing ages. The time spent in these subclasses are what the holding time function emulates in a semi-Markov process.

A similar approach was taken by Pastor et al. (2005), who modeled the response to harvesting of two classes of forests in northern Minnesota, “regeneration” (young forests up to 20 years old originating after a large canopy clearing such as after a clearcut or catastrophic fire) and “forest” (stands 21 years to 150 years old). The regeneration and forest classes were subdivided into annual age classes. This resulted in a 150×150 matrix:

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & \cdots & p_{1,150} = 1 \\ 1 - p_{11} & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 - p_{12} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 - p_{13} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 - p_{1,149} & 0 \end{bmatrix}$$

Note the similarity to a Leslie Matrix in that the first row and subdiagonal are nonzero but the remaining entries are zero. The first row is the probability of any age class being converted to a 1-year forest after a clearcut. It is assumed that any stand that reaches 150 years will be clearcut, hence $p_{1,150} = 1$. The remaining probabilities in row 1 are calculated according to the following function:

$$p_{1i}(t) = Ae^{-\frac{(t-t_0)^2}{2\sigma^2}}$$

The parameters A , t_0 , and σ represent the maximal percentage of land in age class i which is harvested; the rotation age, or the expected age when the probability of harvest is maximal; and the spread of the distribution of cutting across age classes greater than or less than the rotation age, respectively. The shape of this function is a Gaussian normal function. Since the model is solved with an annual time step, what is not harvested is then moved to the next older age class after each time step.

Pastor et al. (2005) then examined the sensitivity of the modeled projections to these three parameters. Not surprisingly, shortening of the rotation age (t_0) increased the proportion of the landscape in regeneration. In addition, increasing the probability of harvest at rotation age increased the proportion of land in the regeneration class, also not surprising. However, the distribution of harvesting around the rotation age (σ) also had a large impact on the long-term dynamics of the forest. As this spread is increased, the landscape becomes more dominated by the regeneration class. The results for the wider spreads of cutting around rotation age are similar to those of a first-order Markov model, where the probability of cutting within a class is independent of age but in which the same acreage of

land is cut each year as in this model. The reason for the convergence of the two models as σ increases is that in the limit as $\sigma \rightarrow \infty$ the probability of being cut at any one age within a class is equal to that at any other age in that class and the model converges to the first-order Markov model. If the spread of distribution of cutting around the rotation age is narrow, then once a stand passes the rotation age the probability of being cut declines rapidly. The stand is then more likely to pass into older ages and has only a low probability of passing into the younger-aged regeneration class. A narrow spread of harvesting around a nominal rotation age therefore had the surprising effect of increasing the eventual proportion of land in older age classes, despite a relatively high probability of a stand being cut at the nominal rotation age.

Metapopulation models

The idea of a transition probability between different classes is also used to derive metapopulation models. A metapopulation is a population consisting of two or more subpopulations in different places and between which individuals can migrate. Organisms travel between these subpopulations with specified probabilities and go extinct in places with other probabilities. Hanski (1999) is an excellent introduction to metapopulation modeling.

We will now derive the most basic of the metapopulation models, known as the Levins model (Levins 1969), which bears some resemblance to the logistic model. The derivation of other metapopulation models follows this format, more or less. The Levins metapopulation model simulates the occupancy of a landscape by a species colonizing some portions of the landscape and going extinct in other portions.

Assume there are H pixels of habitat for an organism in a landscape. The organism occupies O pixels of habitat at any time t (we will model the rate of change in O). The number of unoccupied pixels at any one time is $H - O$. In some occupied pixels, the organism goes extinct and the pixel becomes unoccupied with probability u , but before it does, it sends out progeny to colonize other unoccupied pixels of habitat with probability c .

What is the change in the number of occupied pixels, and therefore the potential growth of the population across the landscape? The rate at which progeny are colonizing unoccupied habitats is:

$$cO(H - O)$$

and the rate at which occupied habitats are becoming unoccupied through local extinction is:

$$uO$$

so the net rate of change of occupied pixels is:

$$\Delta O = cO(H - O) - uO$$

Now this looks suspiciously like the logistic equation. Can we express the metapopulation dynamics in terms of r and K ?

First factor out an O :

$$\Delta O = [c(H - O) - u]O$$

Then after multiplying through the parentheses by c , rearrange to obtain:

$$\Delta O = [(cH - u) - cO]O$$

Now, this is in the form of the starting point for the logistic equation:

$$\Delta N = [(B_0 - D_0) - (b + d)N]N$$

where

$$r = B_0 - D_0$$

and

$$K = \frac{r}{b + d}$$

only here

$$r = cH - u$$

and

$$K = \frac{cH - u}{c}$$

In this case, K is the proportion of H that is occupied at equilibrium. Therefore, in the limit as $\Delta t \rightarrow dt$

$$\frac{dO}{dt} = (cH - u)O \left(1 - \frac{cO}{cH - u} \right)$$

Multiplying through by $(cH - u)$ yields the final Levins metapopulation model:

$$\frac{dO}{dt} = O \left[(cH - u) - \frac{(cH - u)cO}{cH - u} \right] = O[(cH - u) - cO] \quad (15.5)$$

For the species to persist, dO/dt must be greater than 0 but this only happens when $cH > u$. In other words, the maximum rate of colonization, cH , must exceed the per capita rate of local extinction, u . But the maximum rate of colonization depends on the amount of habitat available, H . If H declines due to climate change, competition with other species, or bad management, then the population increase declines as $cH \rightarrow u$, and the population size declines when $cH < u$.

Exercise 15.5

Our friend the snail on a rock is a member of a larger population, but each snail occupies only one rock at a time and each rock can support at most only one snail. The particular pool that the population occupies contains 1000 rocks of which 800 are suitable habitat. Snails leave rocks to find more food or send progeny out into the cruel world to colonize unoccupied rocks; most of the snails and their children die cruel horrible deaths when they are eaten by sculpins. The probability that a snail on a rock colonizes or sends out a progeny to successfully colonize an unoccupied rock is very low, only 0.01 per year. The probability that a snail on a rock will be eaten by a sculpin is quite high, 0.8 per year.

1 Given the Levins population model, can the snail population persist in this environment?

2 Only rocks with algae are suitable habitat because snails eat algae. The climate starts warming and the higher stream temperatures cause the number of rocks with algae to decline exponentially at a rate of 0.1 per year. How many years will it take before the snail population as a whole becomes unsustainable?

Summary: what have we learned?

By defining a landscape as a bounded surface of elements assigned to a finite number of classes, we can conceive of its dynamics as transitions between those classes that take place with certain probabilities. These probabilities can be fixed or they can be functions of various environmental factors, including the state of the landscape as a whole and the proportional distribution of classes it contains. The dynamics of the landscape can therefore be modeled using either transition matrices or differential equations, where the elements of the matrix or the parameters of the differential equations are probabilities of changing from one class to another. While the data used to parameterize these landscape models is often spatially referenced (air photos, remotely sensed images), space is implicit in the models rather than explicit. The advantage of such models is that the same techniques for analyzing population matrix models (Leslie or life-stage models) or ordinary differential equation models of populations and ecosystems can be applied to these landscape models. The disadvantage is that the predictions are aggregate predictions for the landscape as a whole, rather than spatially explicit descriptions of the dynamics. Obtaining spatially explicit predictions requires indexing the population or resource in both time and space, which is considerably more difficult and which will be taken up in the next chapter.

Open questions and loose ends

Further progress in developing transition models can be made by replacing the transition probabilities with functions that are biologically interpretable, such as the function for the probability of harvesting around a nominal rotation age discussed above. Such semi-Markov matrices could help show the consequences of different land management decisions on future landscapes. As with Leslie matrices, transition probability models of landscapes can be useful tools for forecasting the consequences of today's policies if they do not change, or of separating out the effect of different aspects of a policy, such as nominal rotation age, harvest intensity, and the spread of harvest probability around the rotation age.

It is also possible to couple transition matrices of different portions of the landscape through their vectors. For example, we have seen that beaver affect the transition matrices of valley bottoms as ponds are flooded and abandoned. Beaver choose to build ponds where they can obtain aspen, their preferred food, in the adjacent upland. Once the aspen is cut, the riparian zone adjacent to the pond often is converted to the unpreferred conifers such as spruce, pine, or fir. The pond is then abandoned. The probability of a pond being in a flooded (occupied)

state is therefore high if it is adjacent to aspen, but aspen adjacent to an occupied pond has a high probability of being converted to coniferous forest, whereupon the beaver abandon the pond and the pond undergoes a transition from flooded pond to moist meadow. The landscape could then be modeled by two transition matrices, one for the valley bottoms occupied by the ponds or meadows, and one for the adjacent uplands occupied by aspen and conifers. Beaver would then be information agents coupling the two matrices as they make decisions of what to do in the uplands and valley bottoms depending on the distribution of classes in the upland and valley bottom vectors. The transition probabilities in the upland matrices could be functions of the amounts or ratios of different pond or meadow states in the vector for the valley bottoms. Likewise, the transition probabilities of the valley bottoms could be functions of the amounts or ratios of different upland forest states. Such coupled transition matrices could exhibit a wide variety of dynamics, from stable distributions of land classes, to something analogous to limit cycles, and perhaps even chaotic dynamics.

16 Diffusion, advection, the spread of populations and resources, and the emergence of spatial patterns

Space is not homogeneous: organisms disperse throughout it and alter populations as they compete or consume other species; water and air flow across the landscape and take resources and organisms with them; fires spread and kill organisms. We need to come to grips with the problem of the spatial spread of organisms and resources by relaxing the assumption that neither population density nor resource density are homogeneous in space. In this chapter we will explicitly consider changes to organism and resource densities in both space and time. We will see that if species interact and have different diffusion rates, spatial patterns can arise even if the diffusion is random over an otherwise homogeneous landscape.

Derivation of the first-order Fickian diffusion equation

A population of organisms dispersing through space is, to a first approximation, the same as a population of gas molecules dispersing through a room. The theory that describes how the concentration of gas or population density of organisms changes at a given point in space through time is diffusion theory. Berg (1993) is an excellent introduction to the mathematics of diffusion in biological systems and Okubo and Levin (2001) is a very readable but more comprehensive treatise.

We start with the simplest case – a population disperses at random and does not affect its environment. Let's simplify space to position along a line, rather than position in a two-dimensional landscape. We will be concerned only with movement along a line, that is, in one dimension only. Random motion along a one-dimensional line requires that the probabilities of stepping to the right or left are both $1/2$. We first wish to know: what is the net flux, ϕ , of a population across a position x midway between $x - \delta$ and $x + \delta$, where δ is some increment in space, during some increment of time, τ . Once we develop an expression for the net flux of the population, we then proceed to develop a model for the rate of change in population density at any point along the line. This will be our model for the diffusion of a population randomly from a point of introduction in space.

First define the number of organisms at $x - \delta$, x , and $x + \delta$ as $n_{x-\delta}$, n_x , and $n_{x+\delta}$ respectively. The changes in population to the right and left of x are, respectively,

$$\Delta \vec{n} = n_x - n_{x-\delta}$$

$$\Delta \tilde{n} = n_x - n_{x+\delta}$$

Note that if $n_x < n_{x+\delta}$, then Δn is negative and by convention there is a net movement to the left.

Now multiply these expressions by δ/δ :

$$\frac{\delta}{\delta}(n_x - n_{x+\delta}) = \delta \left(\frac{n_x - n_{x+\delta}}{\delta} \right) = \delta(N_x - N_{x+\delta})$$

$$\frac{\delta}{\delta}(n_x - n_{x-\delta}) = \delta \left(\frac{n_x - n_{x-\delta}}{\delta} \right) = \delta(N_x - N_{x-\delta})$$

which gives us the average change in population density ($\Delta N = \Delta n/\delta$) within a distance δ to either side of x .

The net rate of change in the population density, or the net flux ϕ , at x over some time interval τ due to opposing random movements to the right and left is the average change in population density times the probability of stepping right or left ($1/2$), times the distance along which the population flow occurs (δ):

$$\vec{\phi} = \delta \bar{p} \frac{1}{\tau} (N_x - N_{x-\delta}) = \frac{\delta(N_x - N_{x-\delta})}{2\tau}$$

$$\vec{\phi} = -\delta \bar{p} \frac{1}{\tau} (N_x - N_{x+\delta}) = -\frac{\delta(N_x - N_{x+\delta})}{2\tau}$$

(again, by convention, movement to the left is negative). Add these two opposing fluxes to get total net flux across x , then multiply again by δ/δ , gather terms, and simplify:

$$\phi = \frac{\delta}{2\tau} \frac{\delta}{\delta} (N_x - N_{x-\delta} - N_x + N_{x+\delta}) = \frac{\delta^2}{2\tau} \frac{\Delta N \Big|_{x-\delta}^{x+\delta}}{\delta}$$

The term in ΔN is the average change in population density within a distance δ to either side of x . Since δ and τ are constants, we can rename $\delta^2/2\tau = D$. D is known as the diffusion coefficient, whose units are distance² per unit time. This gives us:

$$\phi = D \frac{\Delta N \Big|_{x-\delta}^{x+\delta}}{\delta}$$

Now, take the limit as $\delta \rightarrow 0$. The distance moved then becomes ∂x and the change in population density becomes ∂N , yielding:

$$\phi_x = D \frac{\partial N}{\partial x}$$

In other words, the net flux across x is proportional to the population density gradient at x , the proportionality constant being D , the diffusion coefficient.

But we wish to know: what is the net change in N per unit time across the distance δ to either side of x . Assuming conservation of matter – that is, organisms are neither added or deleted from the population – then the number of individuals

passing x over time τ from the left is $\phi_{x-\delta,x}/\delta$ and the number leaving to the right is $\phi_{x,x+\delta}/\delta$. The net change in population density in the region $x \pm \delta$ is

$$\frac{(N_{x+\delta} - N_x)}{\tau} = \frac{(\phi_{x,x+\delta} - \phi_{x-\delta,x})}{\delta}$$

Again, taking the limit as $\tau \rightarrow 0$ and $\delta \rightarrow 0$, this becomes:

$$\frac{\partial N}{\partial t} = \frac{\partial \phi}{\partial x}$$

Finally, by substituting

$$D \frac{\partial N}{\partial x}$$

for ϕ , we obtain the first-order Fickian diffusion equation:

$$\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2} \quad (16.1)$$

In other words, the change in population density in an infinitesimal region of space at a point x at time t , $\partial N/\partial t$, is proportional to the change in the population density gradient across that point, $\partial^2 N/\partial x^2$, the proportionality coefficient being the diffusion coefficient D .

What does Eq. 16.1 remind you of? It looks like Eq. 3.2, the linear first order differential equation whose solution is exponential growth or decay:

$$\frac{dN}{dt} = rN$$

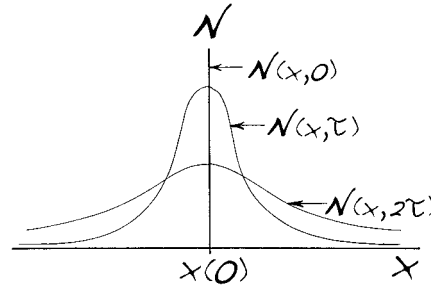
only here the change in N with respect to time is not proportional simply to N , but to some other function of N , namely the second derivative with respect to space. The diffusion coefficient, D , plays the same role as the growth or decay coefficient, r . This is how spatial position affects the time dynamics of a population at a point in space.

Because of the formal similarity of Fickian diffusion (Eq. 16.1) to exponential growth (Eq. 3.2), it should come as no surprise that the solution to the Fickian diffusion equation contains an exponential function whose rate of change is determined by D . However, the solution depends on the initial (boundary) conditions. Consider the case where a population with initial population density N_0 is released at point $x = 0$ at $t = 0$. How does it spread out over time and space? Here's the solution:

$$N_{x,t} = \frac{N_0}{\sqrt{4\pi Dt}} e^{-\frac{1}{4D}\left(\frac{x^2}{t}\right)} \quad (16.2)$$

and Fig. 16-1 shows how the solution evolves in time and space. The biologist J. B. S. Haldane once said that the plain English interpretation of this equation is: "It oozes." You can see from Fig. 16-1 why this is so.

Fig. 16-1 Time evolution of a diffusing population introduced at $x = 0$ at $t = 0$. Shown are solutions at different time steps, τ .



The equation can be extended into more dimensions by simply adding more partial derivatives with respect to each dimension:

$$\frac{\partial N}{\partial t} = D \left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} + \frac{\partial^2 N}{\partial z^2} \right) = D \nabla^2 N$$

where

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$$

is known as the Laplacian gradient operator (in this case applied to N) and is computed by taking the second derivatives of N with respect to each spatial dimension and adding them up. This assumes that the change in N with respect to any direction is independent of any other direction. This is true if the movement is random, but not otherwise. In cases where movement in one direction depends on the gradient of population density in the other dimensions (which have not yet been considered often in ecology), the Laplacian gradient operator is replaced by a matrix known as a tensor matrix.

A process related to diffusion but even more passive is advection. Advection is the passive transport of materials or organisms by the flow of a fluid in which they are embedded (air or water). The movement arises not from an organism or molecule itself moving to the right or left, but by being carried along by the moving fluid, such as a current. In advective flow, the rate of change of population density or nutrient concentration is proportional not to the second partial derivative with respect to space, but to the first partial derivative, $\partial N / \partial x$. Thus advective flow is given by the equation:

$$\frac{\partial N}{\partial t} = A \frac{\partial N}{\partial x}$$

Exercise 16.1

Kareiva (1982) found that the flea beetle *Phyllotreta crucifera*, which feeds on collards, had a diffusion coefficient of $38.3 \text{ m}^2/\text{day}$ in a garden in which the collard patches were 6 m apart. If 100 beetles were introduced into a single patch, what would be the expected density in the nearest patch 10 days later?

Biased diffusion

Many microorganisms and animals do not travel at random, but instead move in the direction of increased abundance of resources or decreased concentration of toxins. In bacteria, this is known as chemotaxis (Keller and Segel 1971, Berg 1993). In the derivation of Fickian diffusion, we assumed that the probability of stepping right or left is $1/2$. When the probability is not $1/2$ but instead is a function of an environmental gradient, then we have biased diffusion (Farnsworth and Beecham 1999).

Let's derive a biased diffusion model of an herbivore that steps with higher probability in the direction of greater food abundance. We would like this model to reduce to Fickian diffusion in the case of no spatial gradients of food or in the case of equal probability of stepping right or left in any direction. We assume a population density of the herbivore of $N(t,x)$. First, we must define a gradient of plant (food) abundance, P , within a distance δ to either side of x as:

$$\frac{\Delta \vec{P}}{\delta} = \frac{P_x - P_{x+\delta}}{\delta}$$

$$\frac{\Delta \tilde{P}}{\delta} = \frac{P_x - P_{x-\delta}}{\delta}$$

The average gradient from $x - \delta$ to $x + \delta$ across the midpoint x is then

$$\left(\frac{\Delta P}{\delta} \right)_x = \frac{P_{x+\delta} - P_{x-\delta}}{2\delta}$$

The probability of an animal moving to the right or left is then

$$\vec{p} = \frac{1}{2} \pm \rho \left(\frac{\Delta P}{\delta} \right)$$

where ρ is the responsiveness of an animal to a gradient of food, P , and $+$ means the animal moves to the right while $-$ means the animal moves to the left. If either ρ or $\Delta P/\delta$ is zero, then this reduces to Fickian diffusion.

Now, as before, calculate the flux of the herbivore population, ϕ , from both the right and from the left:

$$\vec{\phi} = \frac{\delta \vec{p} (N_x - N_{x-\delta})}{\tau}$$

$$\tilde{\phi} = - \frac{\delta \tilde{p} (N_x - N_{x+\delta})}{\tau}$$

Only now the probability p of stepping right or left is not $1/2$, but the above expression. As before, add them together and multiply numerator and denominator by δ/δ :

$$\phi = \frac{\delta^2}{2\tau} \left[\left(1 + \rho \frac{\Delta P}{\delta} \right) \frac{\Delta N}{\delta} - \left(1 - \rho \frac{\Delta P}{\delta} \right) \frac{\Delta N}{\delta} \right]$$

As before, define $D = \delta^2/2\tau$ and take the limit in x as $\delta \rightarrow 0$

$$\phi = D \left(2\rho \frac{\partial P}{\partial x} + 1 \right) \frac{\partial N}{\partial x}$$

Also, as before, we saw that

$$\frac{\partial N}{\partial t} = \frac{\partial \phi}{\partial x}$$

So substituting the new expression for ϕ :

$$\frac{\partial N}{\partial t} = D \frac{\partial}{\partial x} \left(2\rho \frac{\partial P}{\partial x} \frac{\partial N}{\partial x} + \frac{\partial N}{\partial x} \right)$$

and applying the chain rule and collecting terms we obtain our biased diffusion equation:

$$\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2} + 2D\rho \frac{\partial P}{\partial x} \frac{\partial^2 N}{\partial x^2} + 2D\rho \frac{\partial^2 P}{\partial x^2} \frac{\partial N}{\partial x} \quad (16.3)$$

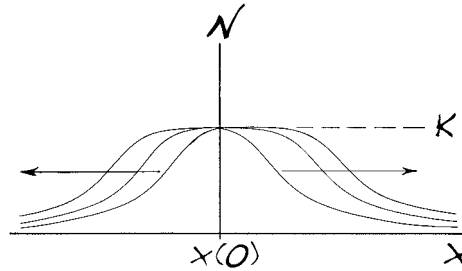
Note that if $\rho = 0$, the herbivore has no response to food gradients and Eq. 16.3 reduces to random Fickian diffusion. Note also that if $\partial P/\partial x = 0$, then there is no gradient in food and the landscape is homogeneous. In this case, the second and third terms vanish and Eq. 16.3 also reduces to Fickian diffusion in a homogeneous landscape. Biased diffusion therefore requires both a gradient in food supply and the ability of the herbivore to respond to it. Because grazing by the herbivore causes a vegetation gradient to form where it has consumed food to where it has not yet consumed food, the second and third terms assume greater importance. The second term modifies the dispersal of the herbivore when the gradient in vegetation is a simple ramp (when $\partial P/\partial x \neq 0$ but $\partial^2 P/\partial x^2 = 0$) and the third term modifies the herbivore dispersal further with more complex patterns.

Therefore, biased diffusion models incorporate more information about the pattern of the landscape (caused in part by consumption of a resource by a consumer) as well as nonrandom behavioral decisions of the consumers. They are derived in the same way as the Fickian (random) diffusion model and in fact contain the expression for Fickian diffusion at their core, just like the logistic model still contains the expression for exponential growth (rN) at its core.

Combining Fickian diffusion with population growth: reaction-diffusion equations

Unless the Fickian diffusion equation is coupled with an equation for population growth at each point x , the population as a whole will approach extinction as it diffuses through space to infinitesimal population densities. Reaction-diffusion equations combine expressions for how population growth responds to the environment along with a diffusion term across the landscape. The “reaction” term can be a logistic growth equation for a one-species system, a set of coupled Lotka–Volterra equations for two-species interactions, or some more complicated model. The diffusion term can

Fig. 16-2 A traveling wave solution to the Fisher equation, Eq. 16.4.



be either simple Fickian diffusion or more complicated nonlinear diffusion models involving feedbacks between D and population density. Holmes et al. (1994) present a survey of representative reaction-diffusion models currently used in mathematical ecology.

The simplest reaction-diffusion model in one dimension is known as the Fisher model (Fisher 1937), and combines logistic growth with Fickian diffusion:

$$\frac{\partial N}{\partial t} = rN \left(1 - \frac{N}{K} \right) + D \frac{\partial^2 N}{\partial x^2} \quad (16.4)$$

Just as simple Fickian diffusion produces a spread of organisms traveling from a population introduction at a particular point, so does the Fisher model. However, this time instead of the peak at the point of introduction declining exponentially with time, the logistic term allows it to maintain itself (eventually at K) through reproduction and population growth at each point. The diffusion term adds immigration and emigration to logistic population growth at each point. The result is a traveling wave rather than an oozing blob (Fig. 16-2). Note that as $t \rightarrow \infty$, $N \rightarrow K$ everywhere and the initial spatial homogeneity caused by the introduction of the population also is destroyed.

The velocity of the wavefront at time t after introduction is given as:

$$v_t = \pm \sqrt{4rD} - b \left(\frac{\ln t}{t} \right) \quad (16.5)$$

where r is the intrinsic rate of population increase from the logistic term, D is the diffusion coefficient, and b is a unit conversion constant determined by the dimensions of space and time. Note that this very quickly converges on the asymptotic velocity $\sqrt{4rD}$ as $t \rightarrow \infty$ because $\ln t/t \rightarrow 0$. In general, the asymptotic velocity of the traveling wavefront in any single species reaction-diffusion model with Fickian diffusion is given by:

$$v_{t \rightarrow \infty} = \sqrt{4f'(0)D}$$

where $f'(0)$ is the eigenvalue of the growth differential equation evaluated at $N = 0$. For example, when the eigenvalue of the logistic equation with respect to N :

$$\frac{df}{dN} = r + 2N \frac{r}{K}$$

is evaluated at $N = 0$, $df/dN = r$ and we recover the first term in Eq. 16.5 for v_i .

What happens if we add diffusion terms to coupled differential equations, such as Lotka–Volterra equations, where each species diffuses at its own diffusion rate? Do each species then reach their equilibrium levels everywhere, as with the Fisher model?

Not necessarily – diffusion can also create and maintain spatial patterns under certain conditions. If the growth of the lowest trophic level involves a positive feedback (autocatalysis) with itself (e.g., population growth) or with some underlying environmental condition (e.g., enhancement of nutrient availability through litter feedbacks), or is sustained by inputs from the surrounding environment, and if the populations of species in the different trophic levels spread or diffuse at different rates, then conditions are ripe for creation of a rich variety of spatial patterns. This phenomenon, known as diffusive instability, was first mathematically described by Turing (1952) and is often called a Turing mechanism in his honor.

To see the conditions under which such spatial heterogeneities arise, consider first a set of coupled equations for the interactions of two species in an otherwise homogeneous environment:

$$\begin{cases} \frac{dN_1}{dt} = f_1(N_1, N_2) \\ \frac{dN_2}{dt} = f_2(N_1, N_2) \end{cases} \quad (16.6)$$

where N_1 and N_2 are prey and predator, respectively, and f_i are the differential equations (e.g., Lotka–Volterra predator–prey equations) describing their growth and interactions. For what follows, it is important to keep in mind that the growth of the “prey” population at the lowermost trophic level is either self-generating by means of autocatalysis, enhanced by positive feedbacks with some underlying environmental variable, or sustained by input from the outside environment. (In much of the literature on reaction diffusion equations, the prey is termed the “activator” because of the positive feedback and the predator is termed the “inhibitor” because it consumes the prey, but I will continue to use the terms “prey” and “predator” in a general sense as pertaining to a wolf eating a moose, a moose eating a plant, or a plant taking up a nutrient.)

Assume there is a spatially uniform (homogeneous) equilibrium in the absence of diffusion where N_1^* and N_2^* represent equilibrium densities of N_1 and N_2 . This equilibrium is spatially homogeneous and stable if small disturbances of size ΔN_i at a particular location decay exponentially when the system is otherwise spatially homogeneous and near equilibrium. Examples of disturbances of size ΔN_i could be harvesting or stocking of a population at a particular place or enhancement of local nutrient availability by fertilization. As we have seen, for a two-species model the rates by which disturbances decay or grow are given by the trace and determinant of the Jacobian matrix J of partial derivatives evaluated at an equilibrium point. Can we extend this to reaction-diffusion equations?

Recall that if the trace of Eq. 16.6 is negative and the determinant is positive, then the perturbations at a point decay exponentially and the system returns to its homogeneous equilibrium state of N_1^* and N_2^* . Now add diffusion terms to each equation (for simplicity, we will consider diffusion in only one direction):

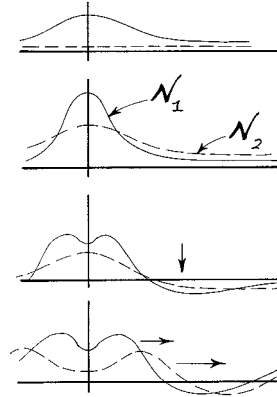
$$\begin{cases} \frac{dN_1}{dt} = f_1(N_1, N_2) + D_1 \frac{\partial^2 N_1}{\partial x^2} \\ \frac{dN_2}{dt} = f_2(N_1, N_2) + D_2 \frac{\partial^2 N_2}{\partial x^2} \end{cases} \quad (16.7)$$

Perturbations to this spatially explicitly model (such as changing the population density of either species, corresponding, for example, to an outbreak, an irruption, stocking, or harvesting) are introduced not simply at a point in time, but at a point in both space and time. Furthermore, the perturbation propagates in space because the diffusion terms “spread” the perturbed populations out in the x direction. The perturbation is further modified by the interactions between the two species which spread or diffuse at different rates.

Before we examine how diffusion changes the Jacobian, let's first take a graphical and intuitive approach to what might happen to the perturbation, as first proposed by Maynard Smith (1971). Let's make the following assumptions, which applies to the predator–prey model with logistic growth of the prey: (i) the growth of the prey (N_1), at least at low population densities, must involve a positive feedback within its own population or with some underlying ecosystem property (exponential growth or logistic growth both fulfill this); (ii) an increase in predator densities (N_2) decreases prey density through consumption; and (iii) the diffusion rate of the predator must be greater than that of the prey. Now imagine that we perturb an otherwise spatially homogeneous population of the prey by stocking it at some arbitrary point, x_0 . This produces a peak in N_1 which, because of assumption (i), then grows. With more food at x_0 , the population of the predator, N_2 , also begins to grow there (Fig. 16-3, top panel). Because of the diffusion term, both the prey and predator populations now start to diffuse outward from x_0 but because of assumption (iii), the spread of the predator outpaces that of the prey. The predator then begins to consume the prey population. Troughs in prey populations develop both at the original peak at x_0 but also partway along the tails of the prey distribution (Fig. 16-3, middle panels). The two split peaks at either side of x_0 now spread outward, followed by the predator populations (Fig. 16-3, bottom panel). Because the predator population always diffuses faster than that of the prey, the predator will always “catch up” with the prey populations and keep forming troughs. Standing waves, traveling waves, and more complicated spatial patterns may now develop, and neither population might decay back to the initial homogeneous equilibrium. We have apparently crossed some critical values of diffusion rates that lead to the initial homogeneous equilibrium becoming unstable and being replaced by a stable spatially patterned set of interacting populations. This is known as a Turing bifurcation, and is our last of the many bifurcations we have seen in this book.

Now let's look at how the diffusion terms enter into the Jacobian. Because the diffusion of either species depends only on itself and not on the other species, the diffusion terms are subtracted from the terms on the main diagonal:

Fig. 16-3 Formation of spatial patterns in a predator–prey model with diffusion in which the population of the predator (N_2) diffuses faster than that of the prey (N_1), modified from figure 38 of Maynard Smith (1971).



$$J_{\text{spatial}} = \left[\begin{array}{cc} \frac{\partial f_1}{\partial N_1} - D_1 \sigma^2 & \frac{\partial f_1}{\partial N_2} \\ \frac{\partial f_2}{\partial N_1} & \frac{\partial f_2}{\partial N_2} - D_2 \sigma^2 \end{array} \right] \Bigg|_{N_1^*, N_2^*, \sigma}$$

where σ is the wavenumber, or the number of the peak in population density assigned in increasing order away from the initial peak that was the perturbation ($\sigma = 1$). σ is proportional to $2\pi/\text{distance between the peaks}$, or the wavelength. The co-existence between N_1 and N_2 is stable and spatially homogeneous when the trace of J_{spatial} is negative and the determinant is positive. The diffusion coefficients obviously make the trace even more negative than in the standard predator–prey equations. The stability of the model to a perturbation therefore depends on what the diffusion terms do to the determinant. If the determinant is negative:

$$\frac{\partial f_1}{\partial N_1} \frac{\partial f_2}{\partial N_2} - \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2} - \sigma^2 \left(D_2 \frac{\partial f_1}{\partial N_1} + D_1 \frac{\partial f_2}{\partial N_2} \right) + \sigma^4 D_1 D_2 < 0 \quad (16.8)$$

then the spatially homogeneous solution will become unstable and the spatially patterned solution will be stable. In what follows, we will assume that $\sigma = 1$, but it is easy to see if we can show that the determinant is negative for this case, then it is negative for all other larger wavenumbers.

Suppose $D_1 = D_2$. Then we can factor out a common D , yielding:

$$\left[\frac{\partial f_1}{\partial N_1} \frac{\partial f_2}{\partial N_2} - \frac{\partial f_2}{\partial N_1} \frac{\partial f_1}{\partial N_2} \right] - D \left(\frac{\partial f_1}{\partial N_1} + \frac{\partial f_2}{\partial N_2} \right) + D^2 < 0$$

Now, the term in brackets is the determinant of the nonspatial predator–prey model without diffusion and the term in parentheses is the trace of the nonspatial model. But if the predator–prey system with diffusion is stable with respect to spatial perturbations, then it is essentially a stable, nonspatial predator–prey model. The determinant of a stable nonspatial model is positive and the trace is negative, so you can see that the above inequality is not possible.

But when $D_2 > D_1$ and D_2/D_1 is greater than some critical value C , then the determinant becomes negative, and diffusive instability sets in. The system becomes spatial heterogeneous as the disturbance propagates across space but does not decrease asymptotically to zero. Under these conditions, the two coexisting species are distributed heterogeneously across the landscape. To see this, we can rearrange Eq. 16.8 to be:

$$\frac{D_2}{D_1} > -\frac{\frac{\partial f_2}{\partial N_2}}{\frac{\partial f_1}{\partial N_1}} + \frac{\det J_{\text{nonspatial}}}{D_1 \frac{\partial f_1}{\partial N_1}} + \frac{D_2}{\frac{\partial f_1}{\partial N_1}}$$

But in predator–prey models, $\partial f_2/\partial N_2$ is negative (recall Chapter 8), so

$$\frac{D_2}{D_1} > \frac{\frac{\partial f_2}{\partial N_2}}{\frac{\partial f_1}{\partial N_1}} + \frac{\det J_{\text{nonspatial}}}{D_1 \frac{\partial f_1}{\partial N_1}} + \frac{D_2}{\frac{\partial f_1}{\partial N_1}}$$

All of these terms are positive. If the predator diffuses much faster than the prey, then D_2/D_1 is very large. In this case, and as Maynard Smith's graphical example (Fig. 16-3) shows, diffusion causes the homogeneous equilibrium to become unstable and spatial patterns to persist, hence the Turing bifurcation. Since the second two terms in the denominator are small, in general if

$$\frac{D_2}{D_1} \gg \frac{\frac{\partial f_2}{\partial N_2}}{\frac{\partial f_1}{\partial N_1}}$$

then spatial homogeneity of two interacting populations of different trophic levels is unstable and spatial heterogeneity of the two interacting populations is stable. In other words, if the ratio of the diffusion rates of the two populations is much greater than their growth rates with respect to themselves, then spatial patterns will spontaneously develop across the landscape.

The most surprising aspect of reaction-diffusion models is that the emergence of spatial patterns is due entirely to the interactions of the two components diffusing randomly at different rates and not necessarily due to any persistent heterogeneity in the underlying environment or preferred directional flow of one or both species. If there are positive feedbacks in the growth of the prey population and greater rates of diffusion of predator than prey, then spatial patterns are almost inevitable. Thus, neither nonrandom foraging of a predator nor underlying environmental gradients are required to produce spatial patterns in predator–prey systems. This is not to say that predators necessarily forage at random nor does it deny the existence of environmental gradients. But, the mere presence of a pattern is not sufficient to invoke them. Put another way, nonrandom foraging and environmental gradients may be sufficient reasons for pattern formation, but they are not necessary reasons.

Unlike predator–prey models, diffusive instability cannot happen in two-species Lotka–Volterra competition models with diffusion (Levin 1974) unless the two

competitors avoid each other (Mimura and Kawasaki 1980), although it can happen in three-species Lotka–Volterra competition models with diffusion (Evans 1980). In addition, diffusive instability and pattern formation can also happen between two competitors and the resource they compete for, which is also a three-dimensional model (Powell and Richerson 1985).

Diffusive instability was first proposed to explain fine-scale spatial heterogeneities of herbivorous zooplankton and phytoplankton in the oceans (Segal and Jackson 1972, Steele 1974, DuBois 1975, Levin and Segal 1976), but later observations showed that patterns of zooplankton and phytoplankton population densities in nature require not only diffusive instability but also directional gradients caused by currents and gyres (Weber et al. 1986). Malchow (2000) is an extensive review of recent developments in the theory of pattern formation in aquatic systems.

Levin (1977) showed that autocatalysis of the prey population is not necessary if the predator consumes prey according to Type II or Michaelis–Menten kinetics. Predator–prey models also apply to a plant taking up a nutrient – the plant is the predator on the nutrient pool. If phytoplankton take up nutrients in a Michaelis–Menten function and herbivores are a constant sink for the phytoplankton, then diffusive instability can occur between phytoplankton and the concentrations of limiting nutrients, resulting in patterned distribution of both plankton and nutrients (Okubo 1978).

Although pattern formation by diffusive instability has been most studied in aquatic systems, there is no reason why it needs to be confined to them. All of the above examples have an herbivore in common. Therefore, there is no reason why pattern formation through diffusive instability cannot arise in terrestrial systems with herbivores, as suggested for tussock moths (Maron and Harrison 1997) and moose (Pastor et al. 1998, 1999b). Terrestrial herbivore populations almost always disperse faster than their forage species disperse seeds, so spatial patterning would seem to be common in those terrestrial ecosystems in which herbivores strongly affect plant community composition and nutrient cycles. If one is working in an ecosystem in which herbivores exert strong control over species composition and/or nutrient cycling rates, then one should immediately suspect diffusive instability as a possible source of any patterns one finds.

Finally, molecules or small organisms in a current are subject to both random movements, captured in the second order diffusive terms, as well as large scale advective flow in a current or sinking out of the water column. A full model of the movement of molecules or small organisms in a current therefore needs to incorporate both diffusion and advection terms as well as inputs from and exports to the external environment and, in the case of organisms, growth regenerative processes. For example, Okubo and Levin (2001, Chapter 3) suggest that the mass balance of nutrients in the sea is then given by:

$$A \frac{\partial N}{\partial x} + D \frac{\partial^2 N}{\partial x^2} + I - lN = 0$$

This is the simplest advection-diffusion model in an open system. Huisman et al. (2006) propose the following reaction-diffusion-advection model for phytoplankton (P) in the vertical (z) direction in the sea:

$$\frac{\partial P}{\partial t} = f(R, I)P - dP - A \frac{\partial P}{\partial z} + D \frac{\partial^2 P}{\partial z^2}$$

where $f(R, I)P$ is the nutrient rate of phytoplankton with respect to nutrient resources, R , and light, I , d is per capita mortality, A is the vertical advective sinking velocity, and D is random diffusion in the vertical direction. This equation is coupled with a reaction-diffusion equation for the nutrient resource:

$$\frac{\partial R}{\partial t} = -f(R, I)P + kdP + D \frac{\partial^2 R}{\partial z^2}$$

where kdP is the rate of regeneration of nutrients by decay of dead phytoplankton. Huisman et al. (2006) show that reducing the diffusive (vertical mixing) term in both equations increases the relative strength of loss of nutrients by advective sinking. This in turn decreases the regeneration rate of nutrients, which in turn destabilizes the system and causes oscillations and possibly chaos in the phytoplankton biomass. Clearly, there is a critical ratio of advection to diffusion that also affects stability.

Summary: what have we learned?

We have relaxed our last assumption in this book, namely that population densities and resource densities are homogeneous in space. Beginning with random dispersal of a population away from a point in space where the population is introduced, we derive first the Fickian diffusion equation. If a population were only governed by this equation without any other growth, population density would return to the initial equilibrium density as the population oozes over the landscape (assuming, of course, the landscape is unbounded). Random diffusion in the absence of any other process destroys inhomogeneities (patterns) in populations over space.

If the landscape has a gradient in resources, caused by some underlying environmental gradient or initiated by consumption of a resource by a population, and if the movement of the consumer is along some gradient of the resource, then we have biased diffusion. We derived an expression for biased diffusion with higher order terms involving the first and second partial derivatives of the resource in space, but which still has the expression for Fickian diffusion at its core.

But patterns need not be caused solely by underlying environmental gradients nor does Fickian diffusion always destroy them. When random Fickian diffusion is coupled with equations for population growth, such as the logistic equation, or equations for the interaction of two populations, such as predator–prey equations, then under certain conditions traveling waves, standing waves, and other patterns may be formed. These models, known as reaction-diffusion models, are particularly prone to pattern formation in the case of predator–prey models and when the predator disperses much faster than the prey. There is a critical ratio of the diffusion rate of the predator to that of the prey above which homogeneous equilibria are no longer stable and the equilibrium becomes a stable standing wave or unstable traveling waves. This critical value is derived in the usual way by examining when the determinant of a spatial Jacobian matrix (containing the diffusion coefficients in addition to the usual elements of population dynamics) becomes negative and hence one of the

eigenvalues becomes positive. We have thus met the last bifurcation in this book, known as a Turing bifurcation in honor of Alan Turing, who discovered it.

Open questions and loose ends

Diffusive instability and spatial pattern formation through trophic interactions is currently an area of theoretical research rich with nontrivial predictions that can be tested experimentally (Okubo and Levin 2001). Some of these experiments may require long-term observations to determine the scales over which spatial patterns arise (e.g., Grünbaum 1992, Pastor et al. 1998) or to determine if spatial patterns change with time. For that matter, it is not yet understood how the length and time scales of patterns related to critical values of model parameters, including not only D but also the parameters governing population growth, interaction, and nutrient flux and regeneration. Perhaps the strongest need in this area right now is for more experimental tests of reaction-advection-diffusion processes that relate the patterns that are formed to the decisions that organisms make when interacting with other species and the surrounding environment. This brings us to the interface between population dynamics, ecosystem ecology, and landscape ecology, an interface which is just beginning to be explored.

How advection and diffusion of organisms vary with age or life stage is also an unexplored area, suggesting the coupling of advection or diffusion models with Leslie or life-stage matrices. Finally, there is a fundamental difference between the macroscopic flux of nutrients and organisms, which is dominated by advection, and the smaller-scale flux of organisms which are controlled by both advection and diffusion to various degrees. The interaction between advection and diffusion on different scales and their respective roles in the stability of populations and ecosystems, as explored by Huisman et al. (2006), deserves much more attention. Is there a critical ratio of advection rate to diffusion rate that destroys patterns, amplifies patterns, or even causes higher order patterns, oscillations, and chaos to emerge? It is entirely possible that a new bifurcation, similar to the Turing bifurcation at a critical value of D_2/D_1 , may arise at some critical value of A/D . This critical value will undoubtedly depend, in turn, on the parameters of the uptake and growth functions which are also included in reaction-advection-diffusion models.

Appendix

MatLab commands for equilibrium and stability analysis of multi-compartment models by solving the Jacobian and its eigenvalues

In the following steps, **boldfaced commands** are **MatLab commands** and *italicized commands* require you to insert a *particular equation* or *numerical value* for a *particular parameter*. The three dots (...) should not be typed, but simply indicate that you should repeat the previous commands for as many equations and parameters as there are in the model. A $\wedge$ means you must type a space at this point.

Step 1. Define the model equations

```
>> dn1 = 'f1(n1, n2,... n_n)'  
>> dn2 = 'f2(n1, n2,... n_n)'  
>> dnn = 'fn(n1, n2,... n_n)'
```

where the functions are, for example, the Lotka–Volterra equations.

Step 2. Find the equilibria

```
>> [n1hat, n2hat, ... nnhat] = solve(dn1,dn2, ... dnn,'n1,n2, ... nn')
```

Step 3. Inspect the equilibria and choose the equilibrium solution of interest

This will be in **n1hat(row,column)** ... **nnhat(row,column)**

Step 4. Solve the Jacobian matrix of partial derivatives

```
>> w = sym('[f1^;f2^; fn]')  
>> v = sym('[n1, n2, ... nn]')  
>> jacob = jacobian(w,v)
```

The column vector **w** is a vector of the equations. You must type in the r.h.s. of the differential equations you defined in Step 1 for f_1 , f_2 , etc., but this can be done by cutting and pasting. Simply typing in **dn1**, etc., won't work. The row vector **v** is a vector of the state variables. The **jacobian** command says: "Find the partial derivatives of the equations in **w** with respect to all the variables in **v** and place them in a matrix called **jacob**." The $(i, j)^{\text{th}}$ entry of the Jacobian is $\partial w_i / \partial v_j$. You can view the general jacobian matrix by typing:

```
>> jacob
```

and you can solve for the general eigenvalues with

```
>> glambdas = eigensys(jacob)
```

but it will be messy. What we are really interested in is the jacobian evaluated at the equilibrium points in **n1hat** and **n2hat**. We must next substitute these for n_1 and n_2 in the jacobian.

Step 5. Substitute the equilibrium values of interest for $n_1 \dots n_n$ found in Step 2 and chosen in Step 3

```
>> jacob = subs(jacob,sym(n1hat(row,column)),^,'n1');
>> jacob = subs(jacob,sym(n2hat(row,column)),^,'n2'); ...
>> jacob = subs(jacob,sym(nnhat(row,column)),^,'nn');
```

Now view the Jacobian evaluated at equilibrium by typing:

```
>> simple(jacob)
```

or

```
>> pretty(jacob)
```

The command **simple** simplifies things algebraically and the **pretty** command makes it look nicer on the screen. You should try both. Both can be combined into one command:

```
>> pretty(simple(jacob))
```

Step 6. Find algebraic expressions for the eigenvalues of the Jacobian matrix evaluated at equilibrium

```
>> elambdas = eigensys(jacob)
```

These symbolic solutions may still be very ugly and almost uninterpretable. In such cases, it may be best to solve the Jacobian and its eigenvalues numerically for particular measured values of the parameters. If you want to numerically evaluate the equilibria, the Jacobian, and the eigenvalues for particular values of parameters, do the following:

Step 7. Assign numerical values to the parameters in the order in which they appear in the equations in Step 1

```
>> parameter 1 = number; parameter 2 = number; ... parameter n = number;
```

where the parameters are, for example, r , K , etc. of Lotka–Volterra equations.

Step 8. Find numerical solutions for the equilibria chosen in Steps 2 and 3 given these parameter values

```
>> n1hat_n = eval(sym(n1hat, row, column))
>> n2hat_n = eval(sym(n2hat, row, column)) ...
>> nnhat_n = eval(sym(nnhat, row, column))
```

Step 9. Find the numerical values of the state variables at equilibrium, the Jacobian evaluated at the equilibria chosen in Steps 2 and 3, and its eigenvalues

```
>> n1 = n1hat_n
>> n2 = n2hat_n ...
>> nn = nnhat_n
>> jac = [eval(diff(dn1, 'n1')) eval(diff(dn1, 'n2')) ... eval(diff(dn1, 'nn'))];
eval(diff(dn2, 'n1')) eval(diff(dn2, 'n2')) ... eval(diff(dn2, 'nn'))); ...
eval(diff(dnn, 'n1')), eval(diff(dnn, 'n2')), ... eval(diff(dnn, 'nn'))]]
>> nlambdas = eig(jac)
```

The command to find the matrix **jac** is an alternative way of solving the Jacobian. Here, the partial derivative of each function with respect to each variable is numerically evaluated separately with **diff**.

Example:

Find the equilibrium points and determine the stability of the coexistence equilibrium for the Lotka–Volterra predator–prey model in Chapter 8:

$$\begin{cases} f_1 = \frac{dN_1}{dt} = N_1(r - hN_2) \\ f_2 = \frac{dN_2}{dt} = N_2(\beta hN_1 - m) \end{cases} \quad (8.2)$$

Step 1. Define the model equations

```
>> dn1 = 'n1*(r-h*n2)';
>> dn2 = 'n2*(b*h*n1-m)';
```

If you don't type a semicolon after each line, MatLab will repeat the equation back to you on the computer screen. This is recommended for proofreading after each equation. If you don't want to see what you typed in, type a semicolon at the end of each line.

*Step 2. Find the equilibria. These will be in the vectors **n1hat** and **n2hat**.*

```
>> [n1hat, n2hat] = solve(dn1,dn2, 'n1,n2')
```

MatLab will return the following:

```
n1hat =
[0]
[ m/b/h]

n2hat =
[0]
[ r/h]
```

You can clean this up with the **pretty** command:

```
>>pretty(n1hat)
>>pretty(n2hat)
```

and MatLab will return

```
[ 0 ]
[   ]
[ m ]
[ --- ]
[b h]
```

for **n1hat** and

```
[ 0 ]
[   ]
[r/h]
```

for **n2hat**. These are each 2×1 vectors. Compare these with Eq. 8.3a,b.

Step 3. Inspect the equilibria and choose the equilibrium solution of interest

This will be in **n1hat(2,1)** and **n2hat(2,1)**, which is the second row, first column of each vector. The first row of each vector contains the trivial solution for n_1 and n_2 .

Step 4. Solve the Jacobian matrix of partial derivatives

First, create a vector **w** of the equations:

```
>> w = sym('[n1*(r-h*n2) ; n2*(b*h*n1-m)]')
```

and MatLab will write the vector back to the screen:

```
w =
[ n1*(r-h*n2)]
[ n2*(b*h*n1-m)]
```

Now create a vector **v** of the variables:

```
>> v = sym('[n1, n2]')
```

and MatLab will write this vector back to the screen:

```
v =  
[ n1, n2]
```

Now solve for the general expression for the Jacobian using the **jacobian** command and name the resulting matrix **jacob**:

```
jacob = jacobian(w,v)
```

MatLab will respond with:

```
[ r-h*n2,    -h*n1]  
[ n2*b*h,  b*h*n1-m]
```

You can clean this up with:

```
>> pretty(jacob)
```

which returns this:

```
  [r - h n2      -h n1  ]  
  [                ]  
  [ n2 b h      b h n1 - m]
```

You can solve the general expressions for the eigenvalues with

```
>> glambdas = eigensys(jacob)
```

but you will get a big mess. Instead, let's insert the equilibrium values for the co-existence solution, which we determined were in **n1hat(2,1)** and **n2hat(2,1)**. We now proceed to Step 5.

Step 5. Substitute the equilibrium values of interest for n1 and n2 found in Step2 and chosen in Step 3

```
>> jacob = subs(jacob,sym(n1hat(2,1)), 'n1');  
>> jacob = subs(jacob,sym(n2hat(2,1)), 'n2');
```

If you want to see what this looks like, type

```
>> pretty(jacob)
```

and MatLab will return the following to the screen:

```
  [ 0  - m/b]  
  [      ]  
  [b r    0 ]
```

Compare this with Eq. 8.6.

Step 6. Find algebraic expressions for the eigenvalues of the Jacobian matrix evaluated at equilibrium

Type

```
>> elambdas = eigensys(jacob)
```

and MatLab will return:

```
elambdas =  
[ (-r*m)^(1/2)  
[ -(-r*m)^(1/2)
```

which you can clean up with

```
>> pretty(elambdas)
```

This returns:

```
[ 1/2 ]  
[ (-r m) ]  
[ ]  
[ 1/2 ]  
[ -(-r m) ]
```

These are the eigenvalues given by Eq. 8.7:

$$\lambda_i = 0 \pm i\sqrt{rm}$$

If you have experimental values for the parameters, you can also solve for numerical values of the equilibria and eigenvalues:

Step 7. Assign numerical values to the parameters in the order in which they appear in the equations in Step 1

```
>> r = 1.5; h = 0.10; b = 0.05; m = 0.05;
```

Step 8. Find numerical solutions for the equilibria chosen in Steps 2 and 3 given these parameter values

```
>> n1hat_n = eval(sym(n1hat, 2,1))  
>> n2hat_n = eval(sym(n2hat, 2,1))
```

and MatLab will return

```
n1hat_n =  
10
```

and


```
n2hat_n =  
15
```

Step 9. Find the numerical values of the state variables at equilibrium, the Jacobian evaluated at the equilibria chosen in Steps 2 and 3, and its eigenvalues

First assign the numerical values of **n1hat_n** and **n2hat_n** to **n1** and **n2**:

```
>> n1 = n1hat_n  
>> n2 = n2hat_n
```

Now evaluate the Jacobian at these numerical values for the parameters and the variables:

```
>> jac = [eval(diff(dn1, 'n1')) eval(diff(dn1, 'n2')); eval(diff(dn2, 'n1'))  
eval(diff(dn2, 'n2'))]
```

MatLab will return the following Jacobian:

```
jac =  
0 -1.0000  
0.0750 0.0000
```

Now calculate the numerical eigenvalues with:

```
>> nlambdas = eig(jac)
```

and MatLab will respond with:

```
nlambdas =  
0.0000 + 0.2739i  
0.0000 - 0.2739i
```

which are imaginary eigenvalues for the Jacobian evaluated at the coexistence equilibrium given the parameter values entered in Step 7.

A good introductory book for MatLab is Pratriap (2005). Roughgarden (1997) is another textbook on theoretical ecology covering some of the material in this textbook but which embeds MatLab commands in the text.

References

Numbers in the margins refer to pages where the reference is cited.

- 211 Aber, J. D., Botkin, D. B., and Melillo, J. M. 1979. Predicting the effects of different harvesting regimes on productivity and yield in northern hardwoods. *Canadian Journal of Forest Research* 9: 10–14.
- 267 Aber, J. D. and Melillo, J. M. 1982. Nitrogen immobilization in decaying hardwood leaf litter as a function of initial nitrogen and lignin content. *Canadian Journal of Botany* 60: 2263–2269.
- 113 Abraham, R. H. and Shaw, C. D. 1988. *Dynamics: The Geometry of Behavior. Part 4: Bifurcation Behavior*. Santa Cruz, California, Aerial Press.
- 278 Acevedo, M. F. 1981. Electrical network simulation of tropical forest succession dynamics. In: Dubois, D. M. (Ed.) *Progress in Electrical Engineering and Mathematical Modelling*, pp. 883–892. Liege, Belgium, Editions CEBEDOC.
- 278 Acevedo, M. F., Urban, D. L., and Abla, M. 1995. Transition and gap models of forest dynamics. *Ecological Applications* 5: 1040–1055.
- 278 Acevedo, M. F., Urban, D. L., and Shugart, H. H. 1996. Models of forest dynamics based on roles of tree species. *Ecological Modelling* 87: 267–284.
- 91 Ackleh, A. S., Marshall, D. F., Heatherly, H. E., and Fitzpatrick, B. G. 1999. Survival of the fittest in a generalized logistic model. *Mathematical Models and Methods in Applied Sciences* 9: 1379–1391.
- 235 Aerts, R. and Heil, G. W. (Eds). 1993. *Heathlands: Patterns and Processes in a Changing Environment*. Dordrecht, The Netherlands, Kluwer Academic Press.
- 211 Ågren, G. I. 1983. Nitrogen productivity of some conifers. *Canadian Journal of Forest Research* 13: 494–500.
- ix, 64 Ågren, G. I. and Bosatta, E. 1998. *Theoretical Ecosystem Ecology: Understanding Element Cycles*. Cambridge, Cambridge University Press.
- 180 Allesina, S. and Pascual, M. 2008. Network structure, predator–prey modules, and stability in food webs. *Theoretical Ecology* 1: 55–64.
- 168, 202, 203, 205, 240 Armstrong, R. A. and McGehee, R. 1980. Competitive exclusion. *The American Naturalist* 115: 151–170.
- 238 Austin, M. P. and Cook, B. G. 1974. Ecosystem stability: a result from an abstract simulation. *Journal of Theoretical Biology* 45: 435–458.
- 166, 167, 174 Ayala, F. J. 1972. Competition between species. *American Scientist* 60: 348–357.
- 166, 174 Ayala, F. J., Gilpin, M. E., and Ehrenfield, J. G. 1973. Competition between species: theoretical models and experimental tests. *Theoretical Population Biology* 4: 331–356.
- 157 Azar, C., Holmberg, J., and Lindgren, K. 1995. Stability analysis of harvesting in a predator–prey model. *Journal of Theoretical Biology* 174: 13–19.
- 278 Baker, W. L. 1989. A review of models of landscape change. *Landscape Ecology* 2: 111–133.
- 11 Barrow, J. 1993. *Pi in the Sky: Counting, Thinking, and Being*. Boston, Back Bay Press.
- 105 Becks, L., Hilker, F. M., Malchow, H., Jürgens, K., and Arndt, H. 2005. Experimental demonstration of chaos in a microbial food web. *Nature* 435: 1226–1229.

- 111 Beddington, J. R. and May, R. M. 1977. Harvesting natural populations in a randomly fluctuating environment. *Science* 197: 463–465.
- 217 Berendse, F. and Aerts, R. 1987. Nitrogen-use efficiency: a biologically meaningful definition? *Functional Ecology* 1: 293–296.
- 57, 59, 63, 236, 245, 267, 268 Berg, B. and McClaugherty, C. A. 2003. *Plant Litter: Decomposition, Humus Formation, Carbon Sequestration*. Berlin, Springer-Verlag.
- 284, 287 Berg, H. C. 1993. *Random Walks in Biology*. Princeton, New Jersey, Princeton University Press.
- 118 Berkes, F., Hughes, T. P., Steneck, R. S., Wilson, J. A., Bellwood, D. R., Crona, B., Folke, C., Gunderson, L. H., Leslie, H. M., Norberg, J., Nyström, M., Olsson, P., Österblom, H., Scheffer, M., and Worm, B. 2006. Globalization, roving bandits, and marine resources. *Science* 311: 1557–1558.
- 105 Berryman, A. A. and Millstein, J. A. 1989. Are ecological systems chaotic – and if not, why not? *Trends in Ecology & Evolution* 4: 26–28.
- 94 Beverton, R. J. H. and Holt, S. J. 1957. *On the Dynamics of Exploited Fish Populations*. London, Her Majesty's Stationary Office, Fisheries Investigations, Series 2, Vol. 19.
- 90 Birch, C. P. D. 1999. A new generalized logistic sigmoid growth equation compared with Richards growth equation. *Annals of Botany* 83: 713–723.
- 237 Bolker, B. M., Pacala, S. W., and Parton, W. J. 1998. Linear analysis of soil decomposition: insights from the Century model. *Ecological Applications* 8: 425–439.
- 73 Bonner, J. T. 1965. *Size and Cycle*. Princeton, New Jersey, Princeton University Press.
- 267, 269 Bosatta, E. and Staaf, H. 1982. The control of nitrogen turn-over in forest litter. *Oikos* 39: 143–151.
- 169 Boucher, D. H. 1985. *The Biology of Mutualism: Ecology and Evolution*. Oxford, Oxford University Press.
- 21 Boyer, C. B. 1959. *The History of the Calculus and its Conceptual Development*. New York, Dover Publications.
- 175 Brauer, F. and Castillo-Chavez, C. 2001. *Mathematical Models in Population Biology and Epidemiology*. Berlin, Springer-Verlag.
- 211–13 Bridgham, S. D., Pastor, J., McClaugherty, C. A., and Richardson, C. J. 1995. Nutrient-use efficiency: a litterfall index, a model, and a test along a nutrient-availability gradient in North Carolina peatlands. *The American Naturalist* 145: 1–21.
- 59 Bridgham, S. D., Updegraff, K., and Pastor, J. 1998. Carbon, nitrogen, and phosphorus mineralization in northern wetlands. *Ecology* 79: 1545–1561.
- 120 Briggs, G. E. and Haldane, J. B. S. 1925. A note on the kinetics of enzyme action. *Biochemical Journal* 19: 339–339.
- 246, 249 Bryant, J. P. and Kuropat, P. J. 1980. Selection of winter forage by subarctic browsing vertebrates: the role of plant chemistry. *Annual Review of Ecology and Systematics* 11: 261–285.
- 209 Burmaster, D. E. 1979. The continuous culture of phytoplankton: mathematical equivalence among three steady-state models. *The American Naturalist* 113: 123–134.
- 121 Burnett, T. 1956. Effects of natural temperatures on oviposition of various numbers of an insect parasite (Hymenoptera, Chalcididae, Tenthredinidae). *Annals of the Entomological Society of America* 49: 55–59.
- 127, 132, 197, 203 Carpenter, S. R. 2003. *Regime Shifts in Lake Ecosystems: Pattern and Variation*. Excellence in Ecology, Vol. 15. Oldendorf/Luhe, Germany, International Ecology Institute.
- ix Case, T. J. 2000. *An Illustrated Guide to Theoretical Ecology*. Oxford, Oxford University Press.
- 5, 7 Caswell, H. A. 1988. Theory and models in ecology: a different perspective. *Ecological Modelling* 43: 33–44.
- 49, 67–71, 73, 76, 77, 277 Caswell, H. 2000. *Matrix Population Models: Construction, Analysis, and Interpretation*, 2nd edition. Sunderland, Massachusetts, Sinauer Associates.
- 236 Chapin, F. S., III, Sydnora Bret-Harte, M., Hobbie, S. E., and Zhong, H. 1996. Plant functional types as predictors of transient responses of arctic vegetation to global change. *Journal of Vegetation Science* 7: 347–358.

- 81 Cisne, J. L. 2005a. How science survived: Medieval manuscripts' "demography" and classic texts' extinction. *Science* 307: 1305–1307.
- 81 Cisne, J. L. 2005b. Response to comment on "How science survived: Medieval manuscripts' 'demography' and classic texts' extinction". *Science* 310: 1618c.
- 249 Cohen, Y., Pastor, J., and Vincent, T. 2000. Nutrient cycling in evolutionary stable ecosystems. *Evolutionary Ecology Research* 6: 719–743.
- 246 Coley, P. D., Bryant, J. P., and Chapin, F. S. III. 1985. Resource availability and plant herbivore defense. *Science* 230: 895–899.
- 105 Constantino, R. F., Desharnis, R. A., Cushing, J. M., and Dennis, B. 1997. Chaotic dynamics in an insect population. *Science* 275: 389–391.
- 24, 49, 213 Courant, R. and John, F. 1965. *Introduction to Calculus and Analysis*, Volume 1. New York, John Wiley and Sons.
- 49, 137 Courant, R. and John, F. 1974. *Introduction to Calculus and Analysis*, Volume 2. New York, John Wiley & Sons.
- 18, 20, 23, 48, 49 Courant, R. and Robbins, H. 1961. *What is Mathematics?* Oxford, Oxford University Press.
- 80 Cramer, J. S. 2003. The origin and development of the logit model. In: Cramer, J. S. (Ed.) *Logist Models from Economics and Other Fields*, Chapter 9. Cambridge University Press, Cambridge, UK.
- 118 Crowe, B. L. 1969. The tragedy of the commons revisited. *Science* 166: 1103–1107.
- 233 Crum, H. 1992. *A Focus on Peatlands and Peat Mosses*. Ann Arbor, Michigan, USA, University of Michigan Press.
- 74 Cullina, W. 2000. *The New England Wildflower Society Guide to Growing and Propagating Wildflowers of the United States and Canada*. New York, Houghton Mifflin & Co.
- 121 Dale, B. W., Adams, L. G., and Bowyer, R. T. 1994. Functional response of wolves preying on barren-ground caribou in a multiple-prey ecosystem. *Journal of Animal Ecology* 63: 644–652.
- 49 Davis, P. J. and Hersh, R. 1981. *The Mathematical Experience*. Boston, Birkhäuser.
- 181 DeAngelis, D. A. 1975. Stability and connectance in food web models. *Ecology* 56: 238–243.
- 238, 239, 244 DeAngelis, D. L. 1992. *Dynamics of Nutrient Cycling and Food Webs*. London, Chapman & Hall.
- 219, 238 DeAngelis, D. L., Mulholland, P. J., Palumbo, A. V., Steinman, A. D., Huston, M. A., and Elwood, J. W. 1989. Nutrient dynamics and food web stability. *Annual Reviews of Ecology and Systematics* 20: 71–95.
- 66, 77 DeAngelis, D. L., Post, W. M., and Travis, C. 1980. *Positive Feedbacks in Natural Systems*. New York, Springer-Verlag.
- 81 Declerq, G. 2005. Comment on "How science survived: Medieval manuscripts' 'demography' and classic texts' extinction". *Science* 310: 1618b.
- 249 de Mazancourt, C. and Loreau, M. 2000. Effect of herbivory and plant species replacement on primary production. *The American Naturalist* 155: 735–754.
- 243, 244 de Mazancourt, C., Loreau, M., and Abbadie, L. 1998. Grazing optimization and nutrient cycling: when do herbivores enhance plant production? *Ecology* 79: 2242–2252.
- 104 Diacu, F. and Holmes, P. 1996. *Celestial Encounters. The Origins of Chaos and Stability*. Princeton, NJ, Princeton University Press.
- 273 Dieckmann, U., Law, R., and Metz, J. A. J. (Eds). 2000. *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge, Cambridge University Press.
- 124 Döbel, H. G. 1987. The role of spiders in the regulation of salt marsh planthopper populations. University of Maryland, Unpubl. Master's Thesis.
- 109 Doebeli, M. 1993. The evolutionary advantage of controlled chaos. *Proceedings of the Royal Society of London B* 254: 281–285.
- 109 Doebeli, M. and Ruxton, G. D. 1997. Controlling spatiotemporal chaos in metapopulations with long-range dispersal. *Bulletin of Mathematical Biology* 59: 497–515.
- 190, 207, 208, 210 Droop, M. R. 1968. Vitamin B₁₂ and marine ecology. IV. The kinetics of uptake, growth, and inhibition in *Monochrysis lutheri*. *Journal of Marine Biological Association, UK* 48: 689–733.

- 261 Droop, M. R. 1974. The nutrient status of algae in continuous culture. *Journal of Marine Biological Association, UK* 54: 825–855.
- 295 DuBois, D. M. 1975. A model of patchiness for prey–predator phytoplankton populations. *Ecological Modelling* 1: 67–80.
- 49 Dunham, W. 1990. *Journey Through Genius: The Great Theorems of Mathematics*. New York, John Wiley & Sons.
- 49 Dunham, W. 2004. *The Calculus Gallery: Masterpieces from Newton to Lebesgue*. Princeton, Princeton University Press.
- 60 Durrett, R. and Levin, S. 1994. The importance of being discrete (and spatial). *Theoretical and Population Biology* 46: 363–394.
- ix Edelstein-Keshet, L. 1988. *Mathematical Models in Biology*. New York, McGraw-Hill.
- 71 Ellner, S. P. and Guckenheimer, J. 2006. *Dynamic Models in Biology*. Princeton, NJ, Princeton University Press.
- 104, 105 Ellner, S. P. and Turchin, P. 1995. Chaos in a noisy world: new methods and evidence from time series analysis. *The American Naturalist* 145: 343–375.
- 138 Elton, C. and Nicholson, M. 1942. The ten year cycle in numbers of lynx in Canada. *Journal of Animal Ecology* 11: 215–244.
- 73, 74 Eriksson, O. 1988. Variation in growth rate in shoot populations of the clonal dwarf shrub *Linnaea borealis*. *Holarctic Ecology* 11: 259–266.
- 73, 74 Eriksson, O. 1992. Population structure and dynamics of the clonal dwarf shrub *Linnaea borealis*. *Journal of Vegetation Science* 3: 61–68.
- 216 Errington, P. L. 1963. *Muskrat Populations*. Ames, Iowa, Iowa State University Press.
- 295 Evans, G. T. 1980. Diffusive structure: counter examples to any explanation. *Journal of Theoretical Biology* 82: 313–315.
- 287 Farnsworth, K. D. and Beechan, J. A. 1999. How do grazers achieve their distribution? A continuum of models from random diffusion to the ideal free distribution using biased random walks. *The American Naturalist* 153: 509–526.
- 82 Feller, W. 1940. On the logistic law of growth and its empirical verifications in biology. *Acta Biotheoretica* 5: 51–66.
- 12 Feynman, R. P., Leighton, R. B., and Sands, M. L. 1963. *Feynman Lectures in Physics*. Reading, Massachusetts, Addison-Wesley.
- 290 Fisher, R. A. 1937. The wave of advance of advantageous genes. *Annals of Eugenics* 7: 355–369.
- 157 Flaaten, O. 1988. *The economics of multispecies harvesting*. Berlin, Springer-Verlag.
- 249 Flanagan, P. W. and Van Cleve, K. 1993. Nutrient cycling in relation to decomposition and organic-matter quality in taiga ecosystems. *Canadian Journal of Forest Research* 13: 795–817.
- 249 Frank, D. A. and Evans, R. D. 1997. Effects of native grazers on N cycling in a north temperate grassland ecosystem: Yellowstone National Park. *Ecology* 78: 2238–2249.
- 155 Freedman, H. I. and Wolkowicz, G. S. K. 1986. Predator–prey systems with group defense: the paradox of enrichment revisited. *Bulletin of Mathematical Biology* 48: 493–508.
- 20 Friedlander, S. 2007. Isaac Newton and Roger Cotes. *Bulletin (New Series) of the American Mathematical Society* 44: 255–257.
- 251 Fry, B. 2006. *Stable Isotope Ecology*. New York, Springer-Verlag.
- 127 Fryxell, J. M., Smith, I. M., and Lynn, D. H. 2005. Evaluation of alternative harvesting strategies using experimental microcosms. *Oikos* 111: 143–149.
- 213 Funk, J. L. and Vitousek, P. M. 2007. Resource-use efficiency and plant invasion in low-resource systems. *Nature* 446: 1079–1081.
- 83 Gabriel, J.-P., Saucy, F., and Bersier, L.-F. 2005. Paradoxes in the logistic equation? *Ecological Modelling* 185: 147–151.
- 49 Gantner, S. 2001. *Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988–1998*. New York, Mathematical Association of America.
- 180 Gardner, M. R. and Ashby, W. R. 1970. Connectance of large, dynamical (cybernetic) systems: critical values for stability. *Nature* 228: 784.

- 261 Gates, C. T. and Wilson, J. R. 1974. The interaction of nitrogen and phosphorus on the growth, nutrient status, and nodulation of *Stylosanthes humilis* H. B. K. (Townsville Stylo). *Plant and Soil* 41: 325–333.
- 82 Getz, W. M. 1996. A hypothesis regarding the abruptness of density-dependence and the growth rate of populations. *Ecology* 77: 2014–2026.
- 111 Getz, W. M. and Haight, R. G. 1989. *Population Harvesting: Demographic Models of Fish, Forest, and Animal Resources*. New Jersey, Princeton University Press.
- 81 Gilman, S. L. and Glaze, F. 2005. “How science survived” – Medieval manuscripts as fossils. *Science* 307: 1208–1209.
- 174 Gilpin, M. E. 1975. Limit cycles in competition communities. *The American Naturalist* 109: 51–60.
- 178 Gilpin, M. E. 1979. Spiral chaos in a predator–prey model. *The American Naturalist* 113: 306–308.
- 81 Gingerich, O. 2005. *The Book Nobody Read: Chasing The Revolutions of Nicholas Copernicus*. New York, Penguin Books.
- 82, 83 Ginzburg, L. R. 1992. Evolutionary consequences of basic growth equations. *Trends in Ecology and Evolution* 7: 133.
- 215 Ginzburg, L. R. 1998. Assuming reproduction to be a function of consumption raises doubts about some population predator–prey models. *Journal of Animal Ecology* 67: 325–327.
- 160 Golubitsky, M. and Stewart, I. 2002. *The Symmetry Perspective*. Boston, Birkhauser.
- 118 Goombridge, B. 1992. *Global Biodiversity: Status of the Earth's Living Resources*. The Netherlands, Chapman and Hall and Kluwer Academic Publishing.
- ix Gotelli, N. 1995. *A Primer of Ecology*. Sunderland, Massachusetts, Sinauer Associates.
- 173 Graves, W., Peckham, B., and Pastor, J. 2006. A bifurcation analysis of a simple differential equations model for mutualism. *Bulletin of Mathematical Ecology* 68: 1851–1872.
- 251 Grenfell, B. T., Heesterbeek, H., Dobson, A. P., Hudson, P. J., and Rizzoli, A. 2002. *The Ecology of Wildlife Diseases*. Oxford University Press, Oxford, UK.
- 73, 76 Griffith, A. B. and Forseth, I. N. 2005. Population matrix models of *Aeschyomene virginica*, a rare annual plant: implications for conservation. *Ecological Applications* 15: 222–233.
- 245 Grime, J. P., Cornelissen, J. H. C., Thompson, K., and Hodgson, J. G. 1996. Evidence of a causal connection between anti-herbivore defence and the decomposition rate of leaves. *Oikos* 77: 489–494.
- 121 Gross, J. E., Shipley, L. A., Hobbs, N. T., Spalinger, D. E., and Wunder, B. A. 1993. Functional response of herbivores in food-concentrated patches: tests of a mechanistic model. *Ecology* 74: 778–791.
- 247 Grover, J. P. 1997. *Resource Competition*. London, Chapman & Hall.
- 297 Grünbaum, D. 1992. Local processes and global patterns: biomathematical models of bryozoan feeding currents and density dependent aggregation in Antarctic krill. PhD Thesis, Cornell University.
- 88 Guckenheimer, J. and Holmes, P. 1983. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York, Springer-Verlag.
- 108, 109 Güemez, J. and Matías, M. A. 1993. Control of chaos in unidimensional maps. *Physical Review Letters* A 181: 29–32.
- 175 Gurney, W. S. C., and Lawton, J. H. 1996. The population dynamics of ecosystem engineers. *Oikos* 76: 273–283.
- 81, 94 Gurney, W. S. C. and Nisbet, R. M. 1998. *Ecological Dynamics*. Oxford, Oxford University Press.
- 36 Gutierrez, A. P. 1996. *Applied Population Ecology: A Supply-Demand Approach*. New York, John Wiley.
- 56 Hall, D. J. 1964. An experimental approach to the dynamics of a natural population of *Daphnia galeata Mendotae*. *Ecology* 45: 94–112.
- 274 Hall, F. G., Botkin, D. B., Strebel, D. E., Woods, K. D., and Goetz, S. J. 1991. Large-scale patterns of forest succession as determined by remote sensing. *Ecology* 72: 628–640.
- 280 Hanski, I. 1999. *Metapopulation Ecology*. Oxford, Oxford University Press.

- 117, 118 Hardin, G. 1968. The tragedy of the commons. *Science* 162: 1243–1248.
- 157 Hardy, A. 1967. *Great Waters*. New York, Harper and Row.
- 167 Harper, J. L. 1979. *Population Biology of Plants*. New York, Academic Press.
- 222 Harwell, M. A., Cropper, W. P., Jr, and Ragsdale, H. L. 1977. Nutrient recycling and stability: a reevaluation. *Ecology* 58: 660–666.
- 156, 178 Hastings, A. and Powell, T. 1991. Chaos in a three-species food chain. *Ecology* 72: 896–903.
- 261 Heady, E. O., Pesek, J. T., and Brown, W. G. 1955. Crop response surfaces and economic options in fertilizer use. *Iowa State College Bulletin* 424, Ames, Iowa, USA.
- 24, 49 Hight, D. W. 1977. *A Concept of Limits*. New York, Dover Publications.
- 111, 118 Hilborn, R., Walters, C. J., and Ludwig, D. 1995. Sustainable exploitation of renewable resources. *Annual Review of Ecology and Systematics* 26: 45–67.
- 101, 104, 131, 150 Hirsch, M. W., Smale, S., and Devaney, R. L. 2004. *Differential Equations, Dynamical Systems, and An Introduction to Chaos*. London, Elsevier Academic Press.
- 236 Hobbie, S. E. 1996. Temperature and plant species control over litter decomposition in Alaskan tundra. *Ecological Monographs* 66: 503–522.
- 249 Hobbs, N. T., Schimel, D. S., Owensby, C. E., and Ojima, D. J. 1991. Fire and grazing in the tallgrass prairie: contingent effects on nitrogen budgets. *Ecology* 72: 1374–1382.
- 278 Hobbs, R. J. and Legg, C. J. 1983. Markov models and initial floristic composition in heathland vegetation dynamics. *Vegetatio* 56: 31–43.
- 44, 276 Hohn, F. E. 1973. *Elementary Matrix Algebra*. New York, MacMillan & Co.
- 114, 121, 124 Holling, C. S. 1959. The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. *Canadian Entomologist* 91: 293–320.
- 290 Holmes, E. E., Lewis, M. A., Banks, J. E., and Veit, R. R. 1994. Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology* 75: 17–29.
- 8 Holton, G. 1978. Analysis and synthesis as methodological themata. In: Holton, G. (Ed.) *The Scientific Imagination. Case Studies*, Chapter 4. Cambridge, Cambridge University Press.
- 11 Holton, G. 1986. Metaphors in science and education. In: Holton, G. (Ed.) *The Advancement of Science, and Its Burdens*, Chapter 12. Cambridge, Cambridge University Press.
- 274 Horn, H. S. 1975. Markovian properties of forest succession. In: Cody, M. L. and Diamond, J. M. (Eds) *Ecology and Evolution of Communities*, pp. 196–211. Cambridge, Massachusetts, Belknap Press of Harvard University Press.
- 296, 297 Huisman, J., Thi, N. N. P., Karl, D. M., and Sommeijer. 2006. Reduced mixing generates oscillations and chaos in the oceanic deep chlorophyll maximum. *Nature* 439: 322–425.
- 231 Huston, M. A. and DeAngelis, D. A. 1994. Competition and coexistence: the effects of resource transport and supply rates. *The American Naturalist* 144: 954–977.
- 106, 108 Hutchinson, G. E. 1948. Circular causal systems in ecology. *Annals of the New York Academy of Science* 50: 221–226.
- 80, 82, 83 Hutchinson, G. E. 1978. *An Introduction to Population Ecology*. New Haven, Yale University Press.
- 278 Johnston, C. A. and Naiman, R. J. 1990. The use of a geographic information system to analyze long-term landscape alteration by beaver. *Landscape Ecology* 4: 5–19.
- 175 Jones, C. G., Lawton, J. H., and Shachak, M. 1994. Organisms as ecosystem engineers. *Oikos* 69: 373–386.
- 106 Jones, G. S. 1962. The existence of periodic solutions of $f(x) = -\alpha f(x-1)[1+f(x)]$. *Journal of Mathematical Analysis and Applications* 5: 435–450.
- 287 Kareiva, P. 1982. Experimental and mathematical analysis of herbivore movement. *Ecological Monographs* 52: 261–282.
- 287 Keller, E. F. and Segal, L. A. 1971. Model for chemotaxis. *Journal of Theoretical Biology* 30: 225–234.
- 190 Kielland, K. 1994. Amino acid absorption by arctic plants: implications for plant nutrition and nitrogen cycling. *Ecology* 75: 2373–2383.
- 82, 130 Kingsland, S. E. 1995. *Modeling Nature*, 2nd edition. Chicago, University of Chicago Press.

- 156 Klebanoff, A. and Hastings, A. 1994. Chaos in a three-species food chain. *Journal of Mathematical Biology* 32: 427–451.
- 13, 49 Klein, F. 1932, reprinted 2004. *Elementary Mathematics from an Advanced Standpoint: Arithmetic, Algebra, Analysis*. Mineola, NY, Dover Publications, Inc.
- 49 Kline, M. 1953. *Mathematics in Western Culture*. Oxford, Oxford University Press.
- 49 Kline, M. 1972. *Mathematical Thought from Ancient to Modern Times*, 3 Volumes. Oxford, Oxford University Press.
- 49 Kline, M. 1986. *Mathematics and the Search for Knowledge*. Oxford, Oxford University Press.
- 210 Knops, J. M. H., Koenig, W. D., and Nash, T. H. III. 1997. On the relationship between nutrient use efficiency and fertility in forest ecosystems. *Oecologia* 110: 550–556.
- 148 Kolomogorov, A. N. 1936. Sulla teoria di Volterra della lotta per l'esistenza. *Giornale Istituto Italiano Attuari* 7: 74–80.
- 80 Kostitzin, V. A. 1937. *Biologie Mathématique*. Paris, Armond Colin.
- ix, 108, 150, 151, 155, 173 Kot, M. 2001. *Elements of Mathematical Ecology*. Cambridge, UK, Cambridge University Press.
- 276 Lancaster, P. and Tismenetsky, M. 1985. *Theory of Matrices: With Applications*, 2nd edition. New York, Academic Press.
- 211 Larcher, W. 1998. *Physiological Plant Ecology*. New York, Springer-Verlag.
- 173 Lee, I. H., Frederickson, A. G., and Tsuchiya, H. M. 1976. Dynamics of mixed cultures of *Lactobacillus plantarum* and *Propionibacterium shermanii*. *Biotechnology and Bioengineering* 18: 513–526.
- 73 Lefkovich, L. P. 1965. The study of population growth in organisms grouped by stages. *Biometrics* 21: 1–18.
- 67 Leslie, P. H. 1945. On the use of matrices in certain population mathematics. *Biometrika* 33: 183–212.
- 295 Levin, S. A. 1974. Dispersion and population interactions. *The American Naturalist* 108: 207–228.
- 295 Levin, S. A. 1977. A more functional response to predator–prey stability. *The American Naturalist* 111: 381–383.
- 295 Levin, S. A. and Segal, L. A. 1976. An hypothesis for the origin of planktonic patchiness. *Nature* 259: 659.
- 280 Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* 15: 237–240.
- 117 Lloyd, W. F. 1833. *Two Lectures on the Checks to Population*. Oxford, Oxford University Press. Reprinted (in part) In: Hardin, G. (Ed.) *Population, Evolution, and Birth Control*. San Francisco, W.H. Freeman, 1964.
- 262, 263, 265, 266, 269 Loladze, I., Kuang, Y., and Elser, J. J. 2000. Stoichiometry in producer-grazer systems: linking energy flow with element cycling. *Bulletin of Mathematical Biology* 62: 1137–1162.
- 244 Loreau, M. 1995. Consumers as maximizers of matter and energy flow in ecosystems. *The American Naturalist* 145: 22–42.
- 231, 232, 243, 249 Loreau, M. 1998. Biodiversity and ecosystem function: a mechanistic model. *Proceedings of the National Academy of Sciences of the United States of America* 95: 5632–5636.
- 104, 156 Lorenz, E. N. 1996. *The Essence of Chaos*. Seattle, Washington, University of Washington Press.
- 108 MacDonald, N. 1978. *Time Lags in Biological Models*. New York, Springer-Verlag.
- 108 MacDonald, N. 1989. *Biological Delay Systems: Linear Stability Theory*. Cambridge, UK, Cambridge University Press.
- 295 Malchow, H. 2000. Non-equilibrium spatio-temporal patterns in models of non-linear plankton dynamics. *Freshwater Biology* 45: 239–251.
- 62 Malthus, T. R. 1826. *An Essay on the Principle of Population*, 6th edition. London, John Murray.
- 295 Maron, J. L. and Harrison, S. 1997. Spatial pattern formation in an insect host–parasitoid system. *Science* 278: 1619–1621.

- 157 Martin, A. and Ruan, S. 2001. Predator–prey models with delay and prey harvesting. *Journal of Mathematical Biology* 43: 247–267.
- 104 Mawhin, J. 2005. Henri Poincaré. A life in science. *Notices of the American Mathematical Society* 52: 1036–1044.
- 148 May, R. M. 1972a. Limit cycles in predator–prey communities. *Science* 177: 900–902.
- 179, 180 May, R. M. 1972b. Will a large complex system be stable? *Nature* 238: 413–414.
- 104 May, R. M. 1974. Biological populations with non-overlapping generations: stable points, stable cycles, and chaos. *Science* 186: 645–647.
- 104 May, R. M. 1975a. Biological populations obeying difference equations: stable points, stable cycles, and chaos. *Journal of Theoretical Biology* 51: 511–524.
- 149, 181, 222, 236 May, R. M. 1975b. *Stability and Complexity in Model Ecosystems*. Princeton, New Jersey, Princeton University Press.
- 104, 105 May, R. M. 1976. Simple mathematical models with very complicated dynamics. *Nature* 261: 459.
- 126 May, R. M. 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* 269: 471–477.
- 98, 100, 106 May, R. M. 1978. Mathematical aspects of the dynamics of animal populations. In: Levin, S. A. (Ed.) *Studies in Mathematical Biology. Part II: Populations and Communities*, pp. 317–366. Washington, DC, Mathematical Association of America.
- 157 May, R. M., Beddington, J. R., Clark, C. W., Holt, S. J., and Laws, R. M. 1979. Management of multispecies fisheries. *Science* 205: 267–277.
- 104 May, R. M. and Oster, G. F. 1976. Bifurcations and dynamic complexity in simple ecological models. *The American Naturalist* 110: 573–599.
- 293 Maynard Smith, J. 1971. *Mathematical Ideas in Biology*. Cambridge, Cambridge University Press.
- 198 Maynard Smith, J. and Price, G. R. 1973. The logic of animal conflict. *Nature* 246: 15–18.
- 109 McCallum, H. I. 1992. Effects of immigration on chaotic population dynamics. *Journal of Theoretical Biology* 154: 277–284.
- 179 McCann, K., Hastings, A., and Huxtel, G. R. 1998. Weak trophic interactions and the balance of nature. *Nature* 395: 794–798.
- 156 McCann, K. and Yodzis, P. 1994. Biological conditions for chaos in a three-species food chain. *Ecology* 75: 561–564.
- 156 McCann, K. and Yodzis, P. 1995. Bifurcation structure of a three-species food chain model. *Theoretical Population Biology* 48: 93–125.
- 180 McMurtrie, A. 1975. Determinants of stability of large, randomly connected systems. *Journal of Theoretical Biology* 50: 1–11.
- 243 McNaughton, S. J. 1979. Grazing as an optimization process: grass-ungulate relationships in the Serengeti. *The American Naturalist* 113: 691–703.
- 238, 249 McNaughton, S. J., Banyikwa, F. F., and McNaughton, M. M. 1997. Promotion of diet-enhancing nutrients by African grazers. *Science* 278: 1798–1800.
- 179 Mehta, M. L. 1967. *Random Matrices*. New York, Academic Press.
- 120 Michaelis, L. and Menten, M. 1913. Die kinetik der invertinwirkung. *Biochemische Zeitung* 49: 333–369.
- 295 Mimura, M. and Kawasaki, K. 1980. Spatial segregation in competitive interaction-diffusion equations. *Journal of Mathematical Biology* 9: 49–64.
- 59 Minderman, G. 1968. Addition, decomposition, and accumulation of organic matter in forests. *Journal of Ecology* 56: 355–362.
- 211 Mitchell, H. L. and Chandler, R. F. 1939. *The Nitrogen Nutrition and Growth of Certain Deciduous Trees of Northeastern United States*. Cornwall on the Hudson, New York, The Black Rock Forest Bulletin 11.
- 120 Monod, J. 1958. *Recherches sur la croissance des cultures bacteriennes*. Paris, Hermann.
- 58, 60 Monsi, M. and Saeki, T. 2005. On the factor light in plant communities and its importance for matter production. *Annals of Botany* 95: 549–567. English translation by M. Schortemeyer

- of the original published as Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion. *Japanese Journal of Botany* 14: 22–52 (1953).
- 218, 220 Morowitz, H. J. 1968. *Energy Flow in Biology: Biological Organization as a Problem in Thermal Physics*. New York, Academic Press.
- 226, 227, 236 Neubert, M. G. and Caswell, H. 1997. Alternatives to resilience for measuring the responses of ecological systems to perturbations. *Ecology* 78: 653–665.
- 49 Neuhauser, C. 2003. *Calculus for Biology and Medicine*, 2nd edition. New Jersey, Prentice-Hall.
- 138 Nicholson, A. J. and Bailey, V. A. 1935. The balance of animal populations. *Proceedings of the Zoological Society of London* 1935: 551–598.
- 222, 238 Nisbet, R. M. and Gurney, W. S. C. 1976. Model of material cycling in a closed ecosystem. *Nature* 264: 633–634.
- 58 Nobel, P. S., Forseth, I. N., and Long, S. P. 1993. Canopy structure and light interception. Pages 79–90. In: Hall, D. O., Scurlock, J. M. O., Bolhar-Nordenkamp, H. R., Leegood, R. C., and Long, S. P. (Eds) *Photosynthesis and Productivity in a Changing Environment: a Field and Laboratory Manual*. London, Chapman and Hall.
- 111 Noy-Meir, I. 1975. Stability of grazing systems: an application of predator–prey graphs. *Journal of Ecology* 63: 459–481.
- 239 O'Brien, W. J. 1974. The dynamics of nutrient limitation of phytoplankton: a model reconsidered. *Ecology* 55: 135–141.
- 141, 242 Oksanen, L. 1983. Trophic exploitation and arctic phytomass patterns. *The American Naturalist* 122: 45–52.
- 141, 242 Oksanen, L. 1988. Ecosystem organization: mutualism and cybernetics or plain Darwinian struggle for existence? *The American Naturalist* 118: 240–261.
- 141, 242 Oksanen, L., Fretwell, S. D., Aruda, J., and Niemela, P. 1981. Exploitation ecosystems in gradients of primary productivity. *The American Naturalist* 131: 424–444.
- 295 Okubo, A. 1978. Horizontal dispersion and critical scales for phytoplankton patches. In: Steele, J. H. (Ed.) *Spatial Pattern in Plankton Communities*, pp. 21–42. New York, Plenum Publishing.
- 284, 296, 297 Okubo, A. and Levin, S. A. 2001. *Diffusion and Ecological Problems*, 2nd edition. New York, Springer-Verlag.
- 57 Olson, J. S. 1963. Energy storage and the balance of producers and decomposers in ecological systems. *Ecology* 44: 321–331.
- 117 Olson, M., Jr. 1965. *The Logic of Collective Association*. Cambridge, MA, Harvard University Press.
- 82 Olson, M. H. 1992. Intuition and the logistic equation. *Trends in Ecology and Evolution* 7: 314.
- 238 O'Neill, R. V. 1976. Ecosystem persistence and heterotrophic regulation. *Ecology* 57: 1244–1253.
- 6, 200, 259, 260–2, 270 O'Neill, R. V., DeAngelis, D. L., Pastor, J., Handley, B. J., and Post, W. M. 1989. Multiple nutrient limitations in ecological processes. *Ecological Modeling* 46: 147–163.
- 108 Ott, E., Grebogi, C., and Yorke, J. A. 1990. Controlling chaos. *Physical Review Letters* 64: 1196–1199.
- 215, 216 Owen-Smith, N. 2005. Incorporating fundamental laws of biology and physics into population ecology: the metaphysiological approach. *Oikos* 111: 611–615.
- 76 Parker, I. M. 2000. Invasion dynamics of *Cytisus scoparius*: a matrix model approach. *Ecological Applications* 10: 726–743.
- 238 Parker, R. A. 1978. Nutrient recycling in closed ecosystem models. *Ecological Modelling* 4: 67–70.
- 108, 109 Parthasarathy, S. and Sinha, S. 1995. Controlling chaos in unidimensional maps using constant feedback. *Physical Review E* 51: 6239–6242.
- 237 Parton, W. J., Stewart, J., and Cole, C. 1988. Dynamics of C, N, P and S in grassland soils: a model. *Biogeochemistry* 5: 109–131.
- 190 Pastor, J. 2003. The mass balances of nutrients in ecosystem theory and experiments: implications for coexistence of species. In: Canham, C. D., Cole, J. J., and Lauenroth, W. K.

- (Eds) *Models in Ecosystem Science*, pp. 272–295. Princeton, NJ, Princeton University Press.
- 278 Pastor, J., Bonde, J., Johnston, C., and Naiman, R. J. 1993a. Markovian analysis of the spatially dependent dynamics of beaver ponds. *Lectures on Mathematics in the Life Sciences* 23: 5–27.
- 6, 212, 217 Pastor, J. and Bridgham, S. D. 1999. Nutrient efficiency along nutrient availability gradients. *Oecologia* 118: 50–58.
- 246, 249 Pastor, J. and Cohen, Y. 1997. Herbivores, the functional diversity of plant species, and the cycling of nutrients in ecosystems. *Theoretical Population Biology* 51: 165–179.
- 295 Pastor, J., Cohen, Y., and Moen, R. 1999b. The generation of spatial patterns in boreal landscapes. *Ecosystems* 2: 439–450.
- 295, 297 Pastor, J., Dewey, B., Moen, R., White, M., Mladenoff, D., and Cohen, Y. 1998. Spatial patterns in the moose-forest-soil ecosystem on Isle Royale, Michigan, USA. *Ecological Applications* 8: 411–424.
- 238, 246 Pastor, J., Dewey, B., Naiman, R. J., McInnes, P. F., and Cohen, Y. 1993b. Moose browsing and soil fertility in the boreal forests of Isle Royale National Park. *Ecology* 74: 467–480.
- 269 Pastor, J. and Durkee Walker, R. 2006. Delays in nutrient cycling and plant population oscillations. *Oikos* 112: 698–705.
- 232, 233 Pastor, J., Peckham, B., Bridgham, S. D., Weltzin, J. F., and Chen, J. 2002. Plant community composition, nutrient cycling, and alternative stable equilibria in peatlands. *The American Naturalist* 160: 553–568.
- 237 Pastor, J. and Post, W. M. 1986. Influence of climate, soil moisture, and succession on forest carbon and nitrogen cycles. *Biogeochemistry* 2: 3–27.
- 279 Pastor, J., Sharp, A., and Wolter, P. 2005. An application of Markov models to the dynamics of Minnesota's forests. *Canadian Journal of Forest Research* 35: 3011–3019.
- 121 Pastor, J., Standke, K., Farnsworth, K., Moen, R., and Cohen, Y. 1999a. Further development of the Spalinger–Hobbs mechanistic foraging model for free-ranging moose. *Canadian Journal of Zoology* 77: 1505–1512.
- 82 Pearl, R. 1925. *The Biology of Population Growth*. New York, Alfred A. Knopf.
- 178 Pesic, P. 2003. *Abel's Proof, An Essay on the Sources and Meaning of Mathematical Unsolvability*. Cambridge, Massachusetts, MIT Press.
- 7 Peters, R. H. 1991. *A Critique of Ecology*. Cambridge, UK, Cambridge University Press.
- 49 Pettofrezzo, A. 1978. *Matrices and Transformations*. New York, Dover Publications.
- 180 Pimm, S. L. 1979. The structure of food webs. *Theoretical Population Biology* 16: 144–158.
- 176, 177, 184 Pimm, S. L. 2002. *Food Webs*. Chicago, University of Chicago Press.
- 106 Pitelka, F. A. 1967. Some characteristics of microtine cycles in the arctic. In: Hansen, H. P. (Ed.) *Arctic Biology*, 2nd edition, pp. 153–184. Corvallis, Oregon, Oregon State University Press.
- 179 Post, W. M., Shugart, H. H., and DeAngelis, D. L. 1978. *Stability Criteria for Multispecies Ecological Communities*. Oak Ridge, Tennessee, USA Oak Ridge National Laboratory Technical Memorandum 6475.
- 295 Powell, T. M. and Richerson, P. J. 1985. Temporal variation, spatial heterogeneity, and competition for resources in plankton systems: a theoretical model. *The American Naturalist* 125: 431–464.
- 304 Pratrapp, R. 2005. *Getting Started with MATLAB 7: A Quick Introduction for Scientists and Engineers*. Oxford, Oxford University Press.
- 181 Quirk, J. P. and Ruppert, R. 1965. Qualitative economics and the stability of equilibrium. *Review of Economics Studies* 32: 311–326.
- 253 Redfield, A. C. 1958. The biological control of chemical factors in the environment. *The American Scientist* 46: 205–221.
- 5 Reiners, W. A. 1986. Complementary models for ecosystems. *The American Naturalist* 127: 59–73.
- 90 Richards, F. J. 1959. A flexible growth function for empirical use. *Journal of Experimental Botany* 10: 290–300.

- 93 Ricker, W. E. 1954. Stock and recruitment. *Journal of the Fisheries Research Board of Canada* 11: 559–623.
- 238, 246 Ritchie, M. E., Tilman, D., and Knops, J. M. H. 1998. Herbivore effects on plant and nitrogen dynamics in oak savanna. *Ecology* 79: 165–177.
- 149, 150 Rosen, R. 1970. *Dynamical Systems Theory in Biology*, Volumr 1. New York, John Wiley and Sons.
- 117 Rosenberg, A. A., Fogarty, M. J., Sissenwine, M. P., Beddington, J. R., and Shepherd, J. G. 1993. Achieving sustainable use of renewable resources. *Science* 262: 828–829.
- xi, 152, 155 Rosenzweig, M. 1971. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. *Science* 171: 385–387.
- 144 Rosenzweig, M. and MacArthur, R. H. 1963. Graphical representation and stability conditions of predator–prey interactions. *The American Naturalist* 97: 209–223.
- ix, 304 Roughgarden, J. 1997. *Primer of Ecological Theory*. Upper Saddle River, NJ, Prentice Hall.
- 105 Schaffer, W. M. 1984. Stretching and folding in lynx fur returns: evidence for a strange attractor in nature? *The American Naturalist* 124: 798–820.
- 178 Schaffer, W. M., Ellner, S., and Kot, M. 1986. Effects of noise on some dynamical models in ecology. *Journal of Mathematical Biology* 24: 479–523.
- 105 Schaffer, W. M. and Kot, M. 1985. Do strange attractors govern ecological systems? *BioScience* 35: 342–350.
- 105 Schaffer, W. M. and Kot, M. 1986. Chaos in ecological systems: the coals that Newcastle forgot? *Trends in Ecology & Evolution* 1: 58–63.
- x, 127, 132 Scheffer, M., Carpenter, S., Foley, J. A., Folke, A., and Walker, B. 2001. Catastrophic shifts in ecosystems. *Nature* 413: 591–596.
- 137 Schey, H. M. 2005. *div, grad, curl, and all that*. London, W. H. Norton and Co.
- 49, 178 Searle, S. R. 1966. *Matrix Algebra for the Biological Sciences*. New York, John Wiley & Sons.
- 262 Seber, G. A. F. and Wild, C. J. 2003. *Nonlinear Regression*, 2nd edition. Wiley Interscience, New York.
- 295 Segal, L. A. and Jackson, J. L. 1972. Dissipative structure: an explanation and an ecological example. *Journal of Theoretical Biology* 37: 545–559.
- 261 Shaver, G. R. and Melillo, J. M. 1984. Nutrient budgets of marsh plants: efficiency concepts and relation to availability. *Ecology* 65: 1491–1510.
- 108, 109 Solé, R. V., Gamarra, J. G., Ginovart, M., and López, D. 1999. Controlling chaos in ecology: from deterministic to individual-based models. *Bulletin of Mathematical Biology* 61: 1187–1207.
- 119, 121 Spalinger, D. E. and Hobbs, N. T. 1992. Mechanisms of foraging in mammalian herbivores: new models of functional response. *The American Naturalist* 140: 325–348.
- 295 Steele, J. H. 1974. Spatial heterogeneity and population stability. *Nature* 248: 83.
- 245 Stephens, D. W. and Krebs, J. R. 1987. *Foraging Theory*. Princeton, NJ, Princeton University Press.
- 190, 208, 211, 252, 254, 269 Sterner, R. W. and Elser, J. J. 2002. *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton, NJ, Princeton University Press.
- 49 Stewart, I. 1992. *The Problems of Mathematics*, 2nd edition. Oxford, Oxford University Press.
- 49 Stewart, I. 1995. *Concepts of Modern Mathematics*. New York, Dover Publications.
- 104 Stewart, I. 2002. *Does God Play Dice? The New Mathematics of Chaos*, 2nd edition. Oxford, UK, Blackwell Publishing.
- 108, 109 Stone, L. 1993. Period-doubling reversals and chaos in simple ecological models. *Nature* 365: 617–620.
- 88, 96, 98, 101, 102, 104, 113, 156 Strogatz, S. H. 1994. *Nonlinear Dynamics and Chaos*. Reading, Massachusetts, Addison-Wesley.
- 200 Taize, L. and Zeiger, E. 2002. *Plant Physiology*. Sunderland, Massachusetts, Sinauer Associates.
- 184, 250 Thébault, E. and Loreau, M. 2003. Food-web constraints on biodiversity–ecosystem functioning relationships. *Proceedings of the National Academy of Sciences of the United States* 100: 14949–14954.

- 255 Tilman, D. 1980. Resources: a graphical-mechanistic approach to competition and predation. *The American Naturalist* 116: 362–393.
- 195, 197, 255 Tilman, D. 1982. *Resource Competition and Community Structure*. Princeton, NJ, Princeton University Press.
- 199 Tilman, D. 1988. *Plant Strategies and the Dynamics and Structure of Plant Communities*. Princeton, NJ, Princeton University Press.
- 273 Tilman, D. and Kareiva, P. (Eds) 1997. *Spatial Ecology*. Princeton, NJ, Princeton University Press.
- 199 Tilman, D. and Pacala, S. 1993. The maintenance of species richness in plant communities. In: Ricklefs, R. E. and Schluter, D. (Eds) *Species Diversity in Ecological Communities: Historical and Geographic Perspectives*, pp. 13–25. Chicago, Illinois, University of Chicago Press.
- 105 Turchin, P. 1993. Chaos and stability in rodent population dynamics: evidence from nonlinear time-series analysis. *Oikos* 68: 167–172.
- 6, 94, 105, 124, 127, 138 Turchin, P. 2003. *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. Princeton NJ, Princeton University Press.
- 291 Turing, A. M. 1952. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London Series B* 237: 37–52.
- 233 Urban, N. R., Eisenreich, S. J., and Bayley, S. E. 1988. The relative importance of denitrification and nitrate assimilation in midcontinental bogs. *Limnology and Oceanography* 33: 1611–1617.
- 249 Van Cleve, K., Oliver, L. K., Schlentner, R., Viereck, L. A., and Dyrness, C. T. 1983. Productivity and nutrient cycling in taiga forest ecosystems. *Canadian Journal of Forest Research* 13: 747–766.
- 169 Vandermeer, J. H. and Boucher, D. 1978. Varieties of mutualistic interactions in population models. *Journal of Theoretical Biology* 74: 549–558.
- 278 van Hulst, R. 1979. On the dynamics of vegetation: Markov chains as models of succession. *Vegetatio* 40: 3–14.
- 80 Verhulst, P. F. 1838. Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique* 10: 113–117.
- 210, 213, 214 Vitousek, P. M. 1982. Nutrient cycling and nutrient use efficiency. *The American Naturalist* 119: 553–572.
- 210, 213, 214 Vitousek, P. M. 1984. Litterfall, nutrient cycling, and nutrient limitation in tropical forests. *Ecology* 65: 285–298.
- 210 Vitousek, P. M. 1997. On regression and residuals: response to Knops et al. *Oecologia* 110: 557–559.
- 223 Vitousek, P. M. and Reiners, W. A. 1975. Ecosystem succession and nutrient retention: a hypothesis. *Bioscience* 25: 376–381.
- 106 Wangersky, P. J. and Cunningham, W. J. 1957. Time lag in prey–predator population models. *Ecology* 38: 136–139.
- 236 Wardle, D. A. 2002. *Communities and Ecosystems: Linking the Aboveground and Belowground Components*. Princeton University Press, Princeton, NJ, USA.
- 295 Weber, L. H., El-Sayed, S. Z., and Hampton, I. 1986. The variance spectra of phytoplankton, krill and water temperature in the Antarctic Ocean south of Africa. *Deep-Sea Research* 33: 1327–1343.
- 222 Webster, J. R., Waide, J. B., and Patten, B. C. 1975. Nutrient recycling and the stability of ecosystems. In: Howell, F. G., Gentry, J. B., and Smith, M. H. (Eds) *Mineral Cycling in Southeastern Ecosystems*, pp. 1–27. Washington, DC, National Technical Information Center.
- 189, 252 Wedin, D. A. 1994. Species, nitrogen and grassland dynamics: the constraints of stuff. In: Jones, C. and Lawton, J. H. (Eds) *Linking Species and Ecosystems*, pp. 253–262. New York, Chapman and Hall.
- 174, 210 Wedin, D. A. and Tilman, D. 1993. Competition among grasses along a nitrogen gradient: initial conditions and mechanisms of competition. *Ecological Monographs* 63: 199–229.

- 184 Wein, R. W. and El-Bayoumi, M. A. 1983. Limitations to predictability of plant succession in northern ecosystems. In: Wein, R. W., Riewe, R. R., and Methven, I. R. (Eds) *Resources and Dynamics of the Boreal Zone*, pp. 214–225. Ottawa, Ontario, Association of Canadian Universities for Northern Studies.
- 245 White, T. C. R. 1993. *The Inadequate Environment*. Berlin, Springer-Verlag.
- 117 Whitehead, A. N. 1948. *Science and the Modern World*. New York, Mentor Books.
- 179 Wigner, E. P. 1959. Statistical properties of real symmetric matrices with many dimensions. In: *Proceedings of the Fourth Canadian Mathematics Conference*, pp. 174–184. Toronto, Canada.
- 10 Wigner, E. P. 1960. On the unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics* 13: 1–14. Reprinted in Wigner, E. P. 1979. *Symmetries and Reflections*. Woodbridge, Connecticut, Ox Bow Press.
- 189, 192 Williams, F. M. 1972. Mathematics of microbial populations, with emphasis on open systems. *Transactions of the Connecticut Academy of Arts and Sciences* 44: 397–426. Reprinted as Deevey, E. S. (Ed.), *Growth By Intussusception; Ecological Essays in Honor of G. Evelyn Hutchinson*, pp. 397–426. Hamden, Connecticut, Archon Books.
- 10 Wilson, E. O. 1998. *Consilience*. New York, Alfred A. Knopf.
- 124 Wise, D. H. 1993. *Spiders in Ecological Webs*. Cambridge, Cambridge University Press.
- 173 Wolin, C. L. and Lowlor, L. R. 1984. Models of facultative mutualism: density effects. *The American Naturalist* 124: 843–862.
- 274 Wolter, P. T. and White, M. A. 2002. Recent forest cover type transitions and landscape structural changes in northeast Minnesota, USA. *Landscape Ecology* 17: 133–155.
- 232 Wright, H. E., Coffin, B. A., and Aaseng, N. E., Editors. 1992. *Patterned Peatlands of Minnesota*. Minneapolis, Minnesota, University of Minnesota Press.
- 180 Yodzis, P. 1980. The connectance of real ecosystems. *Nature* 284: 544–545.
- 179 Yodzis, P. 1981. The stability of real ecosystems. *Nature* 289: 674–676.
- 155, 168 Yodzis, P. 1989. *Introduction to Theoretical Ecology*. New York, Harper and Row.
- 157 Yodzis, P. 1994. Predator–prey theory and management of multispecies fisheries. *Ecological Applications* 4: 51–58.
- 217 Yuan, Z. Y., Li, L. H., Han, X. G., Jiang, F. H., Zhao, M. X., and Lin, G. H. 2004. Effects of plant sizes on the nitrogen use strategy in an annual herb, *Helianthus annuus* (Sunflower). *Acta Botanica Sinica* 46: 889–895.
- 90 Zeide, B. 1993. Analysis of growth equations. *Forest Science* 39: 594–616.
- 156, 157 Zhang, X-an, Chen, L., and Neumann, A. U. 2000. The stage-structured predator–prey model and optimal harvesting policy. *Mathematical Biosciences* 168: 201–210.
- 211 Ågren, G. I. 1983. Nitrogen productivity of some conifers. *Canadian Journal of Forest Research* 13: 494–500.
- 64 Ågren, G. I. and Bosatta, E. 1998. *Theoretical Ecosystem Ecology: Understanding Element Cycles*. Cambridge, Cambridge University Press.

Index

Note: Page numbers in *italics* refer to figures; those in **bold** refer to tables.

- Abel, Niels Henrik, 178
- absorbing classes, 277
- activators, 291
- addition, 12
- Additive model, 260, 261
- advection, 287, 296, 297
- advective flow, 287, 295–6
- age-0 age class *see* young-of-the-year age class
- age-structured models, 65–72, 278–80
 - negative feedbacks, 66
 - positive feedbacks, 66
 - see also* Leslie matrix model
- algae, C:nutrient ratios, 262
- analogy, 11
- aperiodic attractors *see* strange attractors
- aquatic ecosystem photosynthesis, 58
- aquatic systems, pattern formation in, 295
- aspen, 282–3
- asymptotically stable equilibria, 38, 41
 - definition, 40
- asymptotically unstable equilibria, 38–9, 39, 41
 - definition, 40
- autocatalysis, 291, 295
- axioms, 5–6

- basin of attraction, 87
- beaver, 278, 282–3
- bee colony, population growth model, 14
- Berkeley, George, 22
- Beverton–Holt model, 94
- biased diffusion, 288–9
- bifurcations, 55
 - analysis of, 55–6
 - cusp-catastrophe, 126, 126, 127
 - flip/period doubling, 98, 128
 - Turing, 292, 294
 - see also* Hopf bifurcations; saddle nodes/points; transcritical bifurcations
- biodiversity, 232
- biomass increment hypothesis, 223
- biomass production, 207–10
 - rate of *see* productivity
- birth and death as discrete events, 60–2
 - see also* geometric models
- bistability, 126, 163, 258
- “blue sky bifurcations” *see* saddle nodes/points
- bogs, 232–5
- boreal forest, 249
 - fast cyclers, 249
 - slow cyclers, 249
- bottom–up control, top–down control vs., 244–5
- boundary equilibria, 162
- Bureau of Animal Population, 67

- calculus, 21
 - teaching of, 49
- carbon compounds
 - labile fraction, 59, 63
 - recalcitrant fraction, 59, 63
- carbon cycle, 220
- carrying capacity (environmental), 82–3, 92, 114, 189
 - meaning, 129–30
 - in predator–prey models, 129–30, 139, 141
- Cauchy, Augustin-Louis, 22
- Cayley, Arthur, 29
- cell quota, 207
- centers, 47
- CENTURY, 237
- chain rule, of differentiation, 25

- chaos *see* deterministic chaos
- chemostat, models for, 194, 255–6
- chemotaxis, 288
- circle
 - in complex plane, 20
 - unit, 69
- circulation, vector, 136–8, 138
- classes
 - absorbing, 277
 - transient, 277
- closed resource pool, dynamics, 190
- closure, 12
- cobweb plots, 96–7, 96, 97, 98
- coexistence of consumer species
 - litter return and, 228–30
 - with Michaelis–Menten resource uptake, 202–7
 - eigenvalues, 203, 206
 - equilibria, 203, 206–7
 - stability, 203–7, 206
 - stable limit cycle, 203–6, 206
 - with proportional resource uptake, 195–9
 - equilibria, 196–7
 - Jacobian matrix, 198
 - stability, 198–9
 - transcritical bifurcations, 196, 198–9
 - resource pool division and, 230–2
- coexistence of plant species, consumer regulation of, 245–50
- colonization, 280–1
- community matrix, 176–7
- community models, 176–85
 - complexity, 180
 - donor-controlled, 180–1
 - qualitative interaction matrices, 182–3
 - stability, 177–83
 - connectance and, 180, 180, 181
 - succession problem, 184
 - see also* food web models; generalized Lotka–Volterra model
- competition, 159–69
 - detection, 174–5
 - harvesting and, 175
 - limit cycle considerations, 166–7
 - and species invasions, 168–9
 - stable coexistence conditions, 166
 - strong, 162
 - three-species, 174
 - weak, 163
 - see also* community models; Lotka–Volterra competition model
- competition coefficients, 162, 167–8
- competition for resources, theory of, 195
- complex conjugates, 19
 - product of, 20
- complex numbers, 18–21
 - absolute value *see* modulus
 - closure, 19
- complexity, 180
- connectance, 180
 - and stability, 180, 180, 181
- conservation laws, 190
- conservation of mass, 190, 207
 - in closed system, 193
- constant of integration, 27
- constants, differentiation, 25
- consumer–plant–resource systems, 238–51
 - generalization to whole food webs, 250
 - harvesting effects, 251
 - one consumer–one plant species, 239–45, 239
 - equilibria, 239–43, 241, 242, 243
 - nullclines, 240–1, 241
 - transcritical bifurcations, 240, 241, 242
 - one consumer–two plant species, 246–9, 246, 248
 - closed system, 248–9
 - equilibria, 247
 - open system, 246–8
 - stable coexistence region, 248, 248
 - wildlife disease effects, 251
- consumer population response, 110
- continuity, 24
- continuous growth and decay
 - homogeneous population, 53–60
 - see also* exponential models
- continuous models
 - discrete models vs., 61–2
 - see also* exponential models
- continuously differentiable functions, 24
- contour maps, 13
- contradiction, proof by, 17–18
- controlling factors, 59
- conversion efficiency, 130, 143, 252
- cosine functions, 20
- Cotes, Roger, 20
- cropping, 119–20
- curl, of vector field, 136–7
- current, model of movement in, 295–6
- cusp-catastrophe, 126, 126, 127
- cycles, nutrient *see* nutrient cycling
- damping ratio, 70, 277
- Darwin, Charles, 62

- decay rates, 57
 - critical, 228, 236
- decomposition models
 - advances in, 64
 - exponential decay, 57–8
 - double, 59, 63
 - generalization, 59
 - homogeneity assumption, 58–9
- delay times, 202, 259–61
- density, of rational numbers, 15–16
- density-dependent population regulation, 79
- dependent variables, 13
- determinants, 31
 - and eigenvalues, 44
- deterministic chaos, 102–5
 - arguments against, 105
 - arguments for, 105
 - control of, 108–9
 - harvesting functions and, 128
 - evidence in nature, 105
 - in predator–prey models, 156
 - separation from random noise, 105
- DeWit replacement series, 167
- dichotomies, false, 6
- differential equations, 24–5
- differentiation, rules of, 25
- diffusion, 284–97
 - biased, 288–9
 - see also* Fickian diffusion; reaction-advection-diffusion models; reaction-diffusion models
- diffusion coefficient, 285, 286
- diffusive instability, 291, 295–6, 297
- dimension, of system, 41
- discrete logistic growth models, 92–109
 - generalized, 108
 - see also* Beverton–Holt model; discrete quadratic model; Ricker model
- discrete models
 - continuous models vs., 61–2
 - see also* discrete logistic growth models; geometric models
- discrete quadratic model, 92–3, 94–107
 - chaotic behavior, 102–5
 - biological mechanisms producing, 106–7
 - cobweb plots, 96–7, 96, 97, 98
 - eigenvalues, 94–5, 99–100, 104
 - equilibria, 94, 96
 - extinction in, 102
 - fractal structure, 101
 - oscillations, 97–101
 - biological mechanisms producing, 106–7
 - stability, 94–5, 98–9, 101
- discriminant, 166
- diseases in wildlife, 251
- distribution of harvesting around rotation
 - age, 279–80, 281
- divergence, of vector field, 136
- division, 15–16
- domain, 13
- domain of attraction, 87
- donor control, 180–1
- Droop model, 208–10
- ecological stability, 198
- ecosystem nutrient budgets, life history
 - characteristics and, 235
- eigenvalues, 38–40
 - complex, 39–40, 68–9
 - raising to power, 68
 - imaginary part, 39
 - of Jacobian matrices, 43, 45
 - real part, 39–40
 - trace and determinant and, 44
- eigenvectors, 43, 44
 - left, 43, 71
 - right, 43, 71
- elasticity, 70–1, 71, 73
- elemental ratios, 253–4
- Elton, Charles, 67
- emigration, 77
- enrichment, paradox of, 152–3, 152, 153, 206
- environmental carrying capacity *see* carrying capacity (environmental)
- equations, 3
- equilibria, 24
 - classification, 47, 47
 - metastable, 40, 40
 - see also* boundary equilibria; interior equilibria; spirals; stable equilibria; unstable
- equilibrium resource density, 190
- equilibrium solutions, 3, 35–6
- Euler, Leonhard, 20
- Euler's method, of order 1, 27
- eutrophication, 197, 207
- evolutionary stable strategy (ESS), 198
- existence and uniqueness theorem, 131, 149
- exotic invasion, 213
- expectations, 34

- exponential functions, 24
 - differentiation, 26
- exponential models, 53–60
 - continuity assumption, 60
 - double exponential decay models, 59
 - conceptual flaw, 63
 - equilibria, 54
 - generalization, 59
 - homogeneity assumption, 58
 - as limit of geometric model, 62
 - stable exponential decay, 56
 - unstable exponential growth, 56
- extinction, local, 280, 281
- fecundity equation, 65
- feedbacks, 4
 - in age-structured models, 66
 - in ecological systems, 106–7
- Fickian diffusion
 - combining with population growth, 289–96
 - see also* reaction-diffusion models
 - first-order equation, 286
 - derivation, 284–7
- Fisher model, 290–1, 290
- fisheries
 - collapse of, 157
 - overharvesting, 117–18
- fixed effort harvesting *see* proportional harvesting
- flip bifurcations *see* period-doubling bifurcations
- flow, of vector, 41–2
- food abundance gradient, 288
- food chains, length of, 141
- food web models, 176–85
 - qualitative interaction matrices, 182, 183
 - stability, 177–83
 - succession problem, 184
 - see also* generalized Lotka–Volterra model
- foraging
 - nonrandom, 295
 - selective, 246–9
- formal equivalences, 193–4
- fractional loss rate, 192
- “functional response”, 110
 - see also* harvest strategy
- functions, 13
- Fundamental Theorem of Algebra, 20, 44
- Fundamental Theorem of Calculus, 26
- generalized Lotka–Volterra model, 176–81
 - eigenvalues, 178–9, 179
 - equilibria, 177
 - Jacobian matrix, 177
 - stability, 177–8
- geometric models, 60–2
 - critical point, 62
 - Leslie models vs., 67
- global stability, 54–5
- gradient, of function of several variables, 26
- greenhouse gasses, 58
- growth rate *see* productivity
- “half life” of decay, 57
- half-saturation constant, 120, 144
- Hamilton, William Rowan, 29
- harvest, definition, 110–11
- harvest strategy, 110–11
- harvester population response *see* consumer population response
- harvesting, 110–28
 - and chaos onset control, 128
 - and competition, 175
 - controlled harvesting strategies, 128
 - distribution around rotation age, 279–80, 281
 - general form for harvested population, 110
 - maximum sustainable yield, 114
 - and mutualism, 175
 - of predator–prey systems, 156–8
 - see also* Michaelis–Menten harvesting; proportional harvesting; quota harvesting; sigmoidal harvesting
- herbivore maximization, of nutrient cycles, 243, 249
- herding defense, 156
- Hermitian matrix, 227
- Hippasus, 17
- Holling Type I curve, 114
- Holling Type II function, 121, 143
- Holling Type III function, 124, 155
- homogeneous populations
 - birth and death as discrete events, 60–2
 - see also* geometric models
 - continuous growth and decay, 53–60
 - see also* exponential models
- Hopf bifurcations, 98
 - in predator–prey models, 151–3, 151, 152, 153
 - with limiting nutrient, 264, 265

- subcritical, 151
- supercritical, 151
- hyperbolic functions, 120
- hyperbolic harvesting *see* Michaelis–Menten harvesting
- hysteresis, 127
- i, 18
 - see also* imaginary numbers
- identity elements, 29, 30
- identity operations, 13
- imaginary numbers, 18
- immigration, 77
- immobilization, 221, 267, 268–9
- independent variables, 13
- inequality conditions, 13
- inflexion points, 40, 41
- inhibitors, 291
- initial conditions, 27
 - sensitivity to, 102–3
- inorganic resources
 - uptake *see* resource uptake
 - why considered, 189–90
- integration, 26
- interaction coefficients, 169, 173
- interior equilibria, 163
- invasions of species *see* species invasions
- invasive species, matrix model
 - investigations, 76
- inverse operations, 14
- irrational numbers, 17–18
- irreducibility, 276
- Jacobian matrices, 42–3
 - eigenvalues of, 43, 45
 - Matlab commands for solving, 298–304
- John Murray (publisher), 62
- kill rate, per capita, 144
- Kolmogorov, Andrei, 148
- Kolmogorov's theorem, 148–50, 203–5
- Kowa, Takahazu Seki, 31
- Kutta, Martin, 28
- labile fraction, 59, 63
- Lambert–Beers Law, 58, 60
- landscapes
 - definition, 273
 - dynamics of, 273
 - see also* Markov chains of landscapes; metapopulation models
 - land management decision consequences, 282
- Laplace expansion, 31
- Laplacian gradient operator, 287
- Lappland voles, 105
- Law of Diminishing returns, 212
- leaching loss, 223
- leaf area index, 60
- Lefkovitch matrix, 73
 - see also* life-stage matrix
- Leibniz, Gottfried, 21
- lemmings, 107
- Leslie, Patrick H., 67
- Leslie matrix models, 67–72
 - applications, 76–7
 - damping ratio, 70
 - with diffusion, 297
 - eigenvalues, 67–9
 - dominant, 70
 - eigenvectors, 67–8, 69
 - dominant, 70
 - left, 71
 - oscillations, 69
 - right, 71
 - elasticity, 70–1, 71
 - geometric models vs., 67
 - problem with, 77
 - rules of population behavior, 68
 - stability of equilibrium, 67–8
 - stable age distribution, 70
- Levins model, 280–1
- lichens, as obligate mutualisms, 169
- Liebig's Law of the Minimum, 259, 261, 263, 270
- life history characteristics, ecosystem
 - nutrient budgets and, 235
- life-stage matrix models, 73–5
 - applications, 76–7
 - with diffusion, 297
 - eigenvalues, dominant, 73
 - eigenvectors, dominant, 73
 - elasticity, 73
 - problem with, 77
- light, “decay” of, 58
- light extinction coefficient, 58
- limits, 21–4, 56
 - definition, 22–4
- linear equations, 34
- linear functions, derivatives of, 25
- linear models, 8
 - as globally stable/unstable, 54–5
- linear operations, 32–3

- LINKAGES, 237
- Linnaea borealis*, 73–5
- dominant eigenvalues, 75, 75
 - life history stages, 74, 74
- Linnaeus, Carl, 73
- litter
- C:N ratio, 266
 - chemistry, 236
- litter–microbe complexes, stoichiometry
- effect on, 266–9
- litter return
- and coexistence of two species, 228–30
 - and decay of plant parts, 237
 - single species systems, 220–5
 - eigenvalues, 224–5
 - equilibria, 221, 222–5, 223, 225
 - Jacobian matrices, 221, 223
 - linear fluxes, 221–2
 - Lotka–Volterra predator–prey models vs., 224
 - Lotka–Volterra uptake, 222–3
 - nullclines, 222, 223–4, 223, 225
 - stability, 222–5, 225
 - transcritical bifurcation, 224
 - vector fields, 224, 225
 - and stability, 220–5
 - and transient responses to perturbations, 226–8
 - see also* nutrient cycling
- local stability, 87
- logistic growth model, 78–91
- applications, 81
 - behavior at critical values of r , 87–9
 - “constrained”, 83
 - derivation, 78–80
 - discrete *see* discrete logistic growth models
 - eigenvalues, 86
 - explicit solution, 83–5
 - forms, 80–1
 - “free”, 83
 - generalized, 90–1
 - harvesting effects on *see* harvesting interpretation, 81–3
 - oscillatory dynamics, 106–7
 - per capita growth rate, 93
 - population growth according to, 85
 - resource uptake and
 - closed system, 193–4
 - open system, 194–5
 - role in theoretical ecology, 81–2
 - stability
 - algebraic analysis, 86–7
 - geometric analysis, 85–6
- Lotka, A. J., 130
- Lotka–Volterra competition model, 160–8
- with diffusion, 295
 - eigenvalues, 165
 - equilibria, 162–4
 - boundary, 162
 - generalized *see* generalized Lotka–Volterra model
 - Jacobian matrix, 164–5
 - at coexistence equilibrium, 165–6
 - nullclines, 160–2, 161, 166
 - saddle node, 163
 - stability
 - algebraic analysis, 164
 - geometric analysis, 162–4
 - symmetry, 160
 - transcritical bifurcation, 166
 - vector fields, 161, 162–4, 163, 164
- Lotka–Volterra equations, 82
- Lotka–Volterra interaction, 115
- Lotka–Volterra mutualism model, 169–73
- equilibria, 170–2, 171, 172
 - Jacobian matrix, 171–2
 - nullclines, 170, 170, 171, 171, 172
 - stability, 171–3, 171, 172
 - vector fields, 171–2, 171, 172
- Lotka–Volterra predator–prey model, 130–9
- circulation, 136–8, 138
 - with diffusion, 291
 - eigenvalues, 134, 135
 - eigenvectors, 134
 - equilibria, 133–6
 - Jacobian matrix, 133–4, 135–6
 - litter return models vs., 224
 - nullclines, 131–3, 132, 133
 - phase space, 131
 - saddle node, 134
 - stability, 134–7, 139, 145–6
 - vector fields, 131–3, 136–8, 138
 - curl, 137
 - divergence, 137
- Lyapunov exponents, 103–4, 105
- and eigenvalues, 104
- Lyapunov stability, 37, 181, 227
- Malthus, Thomas, 62
- Markov chains of landscapes, 273–8
- see also* semi-Markov models

- Markov matrices, 274
 - coupled transition matrices, 282–3
 - dominant eigenvalue, 276
 - dominant eigenvector, 276
- Markov processes
 - first-order, 274
 - see also* semi-Markov models
- mass balance, of single inorganic resource, 190–2
- material cycling theorem, 218
- mathematical ecology, overview, 3–5
- mathematical models
 - errors of perception of, 7
 - requirements, 7–10
- mathematics
 - abstractness, 48–9
 - history, 49
 - origin, 11
- MathWorld website, 49
- Matlab, commands for equilibrium and stability analysis, 298–304
- matrices, 29
 - addition, 29
 - identity elements, 29, 30
 - irreducible, 276
 - multiplication, 29–30
 - multiplicative inverse, 32
 - operations as linear, 33
 - primitive, 276
 - qualitatively stable, 181
 - reducible, 276
 - square, 30–1
 - subtraction, 29
 - tensor, 287
 - transposition, 31
 - see also* Jacobian matrices; Markov matrices; symmetric matrices
- matrix population models *see* Leslie matrix models; life-stage matrix models
- maximum sustainable yield, 114
- metapopulation, definition, 280
- metapopulation models, 280–1
- Michaelis–Menten functions
 - predation, 120–1, 121, 143
 - resource uptake, 200–2, 259–60
- Michaelis–Menten harvesting, 119–23
 - eigenvalues, 121–2
 - equilibria, 121–3, 122, 123
 - per capita harvest function, 120
 - in predator–prey model *see* Rosenzweig–MacArthur model
 - saddle-node bifurcation, 123, 123
 - stability
 - algebraic analysis, 121–2
 - geometric analysis, 122–3, 123
- Michaelis–Menten production, 207, 208–10, 211
- Michaelis–Menten uptake, 199–202, 209
 - per capita uptake equation, 202
- microbes
 - C:N ratio, 266
 - see also* litter–microbe complexes
- microtines, 107
- min (...) function, 270
- minors, 31
 - expansion by, 31
- modulus, 20
- monotonic laws, 13
- moose, 295
- multiplication, 12
- multispecies community models
 - see* community models
- mutualism, 88, 159, 169–73, 215
 - detection, 174–5
 - facultative, 169, 170, 171
 - facultative–obligate, 172–3
 - harvesting and, 175
 - obligate, 169, 170–1
 - positive feedback limitation, 173
 - strong, 170
 - weak, 170
 - see also* community models; Lotka–Volterra mutualism model
- natural numbers, 12
- natural selection, 62, 91
- net carbon fixation, 210
- net flux, 198, 284
- net survival rate, 81
- neutral stability, 40, 47, 136
- Newton, Isaac, 21
- nitrogen dynamics in decaying litter, 267–9
 - see also* immobilization
- nitrogen productivity, 211
- nodes
 - saddle *see* saddle nodes/points
 - stable, 45
 - unstable, 45
- nonlinear equations, 34
- nonlinear models
 - of single populations, 78–91
 - eigenvalues, 87

- local stability of equilibria, 87
 - see also* logistic growth model
- nullclines, 131
- numbers, 11–12
- “numerical response”, 110
 - see also* consumer population response
- nutrient concentration, 202
- nutrient cycling
 - consumer regulation of *see*
 - consumer–plant–resource systems
 - existence of cycles, 218–20
 - herbivore maximization, 243, 249
 - multiple species with resource pool
 - division, 230–2, 231
 - equilibria, 232
 - stability, 232
 - peatland model, 232–5, 233, 235
 - eigenvalues, 234
 - equilibria, 234–5
 - stability, 234–5
 - transcritical bifurcation, 234–5, 235
 - two-species model with litter return,
 - 228–30, 229
 - eigenvalues, 229–30
 - equilibria, 229
 - stability, 229–30
 - transcritical bifurcations, 229, 230
 - see also* litter return
- nutrient residence time, 216–17
- nutrient turnover rate, 192
- ombrotrophic peatlands, 232–5
- omnivores, 183, 184, 251
- operations, 3, 12–13
- orbit of period 2, 97
- paradox of enrichment, 152–3, 152, 153, 206
- parameters, 3
- partial derivatives, 26
- peatlands
 - ombrotrophic, 232–5
 - pocosin, 213–14
- per capita harvest strategy *see* harvest strategy
- period-2 cycles, 97
- period doubling, 101
 - reversal of, 108–9
- period-doubling bifurcations, 98, 128
- period τ cycles, 97–8, 101
- Perron–Frobenius Theorem, 276
- phase plane, transformations in, 253–5, 253
 - phosphorus dynamics, in decaying litter, 268
 - photosynthesis, 219–20
 - physics, 8
 - phytoplankton
 - in chemostat, 256
 - population density patterns, 295
 - reaction-diffusion-advection model, 296
 - phytoplankton–consumer pelagic system, 239–45
 - pixels, 273, 274
 - plant canopy photosynthesis, 58
 - pocosin peatlands, 213–14
 - Poincaré, Henri, 104
 - Poincaré–Bendixson Theorem, 150, 156
 - Poisson Arrival Time model, 261
 - Poisson process, 260
 - pollutants, introduction of, 118
 - polynomial equations, 20
 - population density gradient, 285
 - population dynamics, of lower plants, 215
 - population regulation, density-dependent, 79
 - power functions, 12
 - derivative, 26
 - negative, 16
 - power rule, of differentiation, 25
 - predator control, 153
 - predator–prey cycles, 138, 138, 139
 - damped, 143
 - predator–prey models, 129–58, 291–5
 - chaos in, 156
 - coupled population oscillations, 154–5
 - with density-dependent self-regulation of prey, 139–43
 - eigenvalues, 140, 142–3
 - equilibria, 140–1, 140, 141
 - Jacobian matrix, 142
 - nullclines, 139–41, 140, 141
 - stability, 140, 141–3, 141
 - transcritical bifurcation, 141, 142, 143
 - vector fields, 139–40, 140, 142
 - with diffusion, 291–5
 - determinant, 293–4
 - Jacobian matrix, 292–3
 - spatial patterns, 293
 - elaborations, 155–6
 - generalized, 148
 - harvesting of predator–prey systems, 156–8
 - Jacobian matrix
 - determinant, 154
 - trace, 154

- with Michaelis–Menten predator harvesting *see* Rosenzweig–MacArthur model
- stability conditions, 154
- stable limit cycles, 148, 150, 151, 152, 167
 - with limiting nutrient, 264
- see also* community models; food web models; Lotka–Volterra predator–prey model
- prediction, 6, 8–9
- primitivity, 276
- processing, 119–20
- product rule, of differentiation, 25
- production of biomass *see* biomass production
- productivity, 207–9, 211–13, 212
- proof by contradiction, 17–18
- proportional harvesting, 114–18, 130
 - eigenvalues, 116
 - equilibria, 115–16, 115
 - implications, 117–18
 - in predator–prey model *see* Lotka–Volterra predator–prey model
- stability
 - algebraic analysis, 115–17
 - geometric analysis, 115, 115, 116, 116
 - transcritical bifurcation, 116–17, 116
- pulsed forcings, 109
- Pythagoreans, 17
- quadratic formula, 44
- qualitatively stable matrices, 181
- Quirk–Ruppert criteria, 181
- quota harvesting, 111–14
 - eigenvalues, 112
 - equilibria, 111–12, 112
 - implications, 117–18
 - perturbations, 112–13
 - saddle-node bifurcations, 113–14, 113, 114
- stability
 - algebraic analysis, 112
 - geometric analysis, 111–12, 112
- quotient rule, of differentiation, 25
- radioactive decay, 57
- range, 13
- ranges of behavior, 5
- rare species, matrix model investigations, 76
- rates of change, 21
 - definition, 16
- rational numbers, 15, 17
 - as dense, 15–16
- reaction–advection–diffusion models, 296, 297
- reaction–diffusion models, 289–96
 - Fisher model, 290–1, 290
 - Lotka–Volterra competition model
 - with diffusion, 295
 - movement in current, 295–6
 - see also* predator–prey models, with diffusion
- reactivity, 227
- recalcitrant fraction, 59, 63
- recursive models, 15
- reducible matrices, 276
- “regime shifts”, 127
- reproductive values, 71–2
- resilient systems, 228
- resistant systems, 228
- resource availability, 211
- resource pool, dividing, 230–2
- resource response curves, crossing of, 211, 216
- resource turnover rate, 192
- resource uptake, 192
 - closed system, 192–4
 - and logistic model, 193–4
 - and coexistence by two consumers *see* coexistence of consumer species
 - open system, 194–5
 - and logistic model, 194–5
- resource use efficiency, 193, 210–14, 212, 213
 - at low resource availability, 213–14
 - nutrient residence time and, 216–17
 - total ecosystem, 213–14
- Ricker model, 93
- right hand rule, 137
- robustness, 6
- Rolle’s Theorem, 213
- roots, 17
 - negative solutions, 18
- Rosenzweig–MacArthur model
 - extension to include limiting nutrient, 262–6
 - equilibria, 264–5
 - Hopf bifurcation, 264, 265
 - nullclines, 264, 265, 266, 266
 - snic bifurcation, 264, 265
 - stable limit cycle, 264
 - transcritical bifurcation, 264, 265

- original, 143–53, 207, 262–3
 - eigenvalues, 151–2
 - equilibria, 144–5, 145, 147–8, 150–2, 151
 - Hopf bifurcations, 151–3, 151, 152, 153
 - Jacobian matrix, 147–8, 151
 - nullclines, 144–8, 151–3, 151
 - stability, 145, 145, 147–8, 150–3, 151
 - stable limit cycles, 148, 150, 151, 152
 - transcritical bifurcation, 145, 151, 151, 152–3, 153
 - vector fields, 144, 146, 148
- rotation age, 279
- Runge, Carl, 28
- Runge–Kutta method, 28
- saddle nodes/points, 46, 113
 - on invariant circle (snic), 264, 265
 - in Lotka–Volterra competition model, 163
 - in Lotka–Volterra predator–prey model, 134
 - in Michaelis–Menten harvesting, 123, 123
 - perturbations around, 46, 46
 - in quota harvesting, 113–14, 113, 114
 - in sigmoidal harvesting, 125–6
- scalars, 29
- self-regulation
 - density-dependent, 179
 - of prey in predator–prey model, 139–43
- semi-Markov models, 278–80
- separatrix, 125, 163
- shade tolerance classification, 211
- sigmoidal harvesting, 123–7
 - equilibria, 124–6, 125
 - implications, 127
 - saddle-node bifurcation, 125–6
 - sigmoidal harvesting function, 124, 155
 - stability
 - algebraic analysis, 125
 - geometric analysis, 125–6, 125, 126
- simulation models, 9, 237
- sine functions, 20
- sink populations, 77, 88
- SIR models, 251
- skeleton, stoichiometry, 270
- source and sink populations, 77
- species invasions, competition and, 168–9
- spirals, 46, 46
 - stable, 47
 - unstable, 47
- stable equilibria, 36–8, 36, 37, 40, 41
 - perturbations around, 45, 45
 - see also* asymptotically stable equilibria
- stable limit cycles, 148
 - existence, 148–50
 - in predator–prey models, 148, 150, 151, 152, 167
 - with limiting nutrient, 264
 - in resource uptake models, 203–6, 206
- stage-structured linear models *see* life-stage matrix models
- state variables, 3
- statistics, 34
 - assumptions, 34
- stoichiometric ratio, 254
- stoichiometry, 252–70
 - allometry of, 270
 - and competition between two species, 255–9
 - equilibria, 256–8, 257, 258
 - nullclines, 256–9, 257
 - stability, 257–8, 258
 - vector field, 256–7, 257
 - constraints on model solutions, 252–3
 - constraints throughout nutrient cycle, 270
 - definition, 252
 - elemental ratios, 253–4
 - importance, 252–3
 - mechanisms of uptake of two resources, 259–62
 - min (...) function, 270
 - open vs. closed systems, 269–70
 - and predator–prey models, 262–6
 - see also* Rosenzweig–MacArthur model
 - and producer–microbial interactions, 266–9
 - transformations in phase plane, 253–5, 253
- strange attractors, 102, 156
 - in three-species models, 178
- structurally unstable equations, 135
- subtraction, 14
- summation rule, of differentiation, 25
- Sylvester, James Joseph, 29
- symbioses, 215
- symmetric matrices, eigenvalues, 178–9, 179
- symmetry, of mathematical object, 160
- Taylor, Brook, 27
- Taylor series, 27, 34
- tensor matrix, 287
- theoretical problems, 5–6
- time delays
 - in ecological systems, 106–7

- in population models, 108
- time-varying solutions, 3
- top-down control, bottom-up control vs., 244–5
- trace, 31, 136
 - and eigenvalues, 44
- “tragedy of the commons”, 117–18
- transcritical bifurcations, 88–9, 89
 - in competition models, 166
 - in nutrient cycling models
 - consumer–plant–resource system, 240, 241, 242
 - peatland, 234–5, 235
 - single species, 224
 - two-species, 229, 230
- in predator–prey models
 - with density-dependent self-regulation of prey, 141, 142, 143
 - with limiting nutrient, 264, 265
 - Rosenzweig–MacArthur model, 145, 151, 151, 152–3, 153
- in proportional harvesting, 116–17, 116
- in resource uptake models, 196, 198–9
- transient classes, 277
- transient responses to perturbations, 226–8
- transition probabilities, 273–4
 - calculation, 274
 - nonstationary behavior, 278
 - normalization, 275–6
- traveling waves, 290
- trigonometric functions, differentiation, 25
- “trivial” solution, 43
- Turing bifurcations, 292, 294
- Turing mechanism, 291
- tussock moths, 295
- twinflower *see* *Linnaea borealis*
- Type I harvesting, 114
- Type II harvesting, 121, 143
- Type III harvesting, 124, 155
- Type IV harvesting, 155–6
- unit circle, 69
- unstable equilibria, 36, 38–9, 39, 40, 41
 - perturbations around, 45, 45
 - see also* asymptotically unstable equilibria
- variables
 - dependent, 13
 - independent, 13
- vector fields, 38, 131
 - circulation, 136–8, 138
 - curl, 136–7
 - divergence, 136
- vector length, species characteristic, 254
- Volterra, Vito, 130–1
- vorticity, 136
- water transparency, 60
- wavenumber, 293
- whales, harvesting of, 157
- wood, stoichiometry, 270
- young-of-the-year age class, 65
 - reproductive value, 71
- zero-net growth isoclines (ZNGIs) *see* nullclines
- zooplankton
 - C:nutrient ratios, 262
 - population density patterns, 295